

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2000, 08:16:18 ; Search time 38.4 Seconds
(without alignments)
244.263 Million cell updates/sec

Title: US-08-900-220-17

Perfect score: 2088
Sequence: 1 MALLNLPLPLCLALLALPA.....PTGMWYSRLYLABELLG 396

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2088	100.0	396	1 W97764	Human Desert hedge
2	2088	100.0	396	1 Y05517	Human Desert hedge
3	2083	99.8	396	1 W97955	Human Desert hedge
4	2008	96.2	396	1 R77345	Human Desert hedge
5	2008	96.2	396	1 W94469	Mouse Dh hedgehog
6	2008	96.2	396	1 W97766	Mouse Desert hedge
7	2008	96.2	396	1 Y05511	Mouse Desert hedge
8	1974	84.5	396	1 R79598	Human Desert hedge
9	1217	58.3	437	1 R77339	Mouse Sonic hedgehog
10	1217	58.3	437	1 W94471	Mouse Sh hedgehog
11	1217	58.3	437	1 W97768	Mouse Sonic hedgehog
12	1217	58.3	437	1 Y05513	Mouse Sonic hedgehog
13	1214	58.1	437	1 W61488	Mouse Sonic hedgehog
14	1198	57.4	425	1 R77338	Chicken Sonic hedgehog
15	1198	57.4	425	1 W94468	Chicken Sh hedgehog
16	1198	57.4	425	1 W97765	Chicken Sonic hedgehog
17	1198	57.4	425	1 Y05510	Chicken Sonic hedgehog
18	1177	56.4	475	1 R77341	Human Sonic hedgehog
19	1177	56.4	475	1 W94473	Human Sh hedgehog
20	1177	56.4	475	1 W97770	Human Sonic hedgehog
21	1177	56.4	475	1 Y05515	Human Sonic hedgehog
22	1173	56.2	462	1 W48736	Human mutated soni
23	1171.5	56.1	411	1 W94474	Human Ih hedgehog
24	1171.5	56.1	411	1 W97763	Human Ih hedgehog
25	1171.5	56.1	411	1 Y05516	Human Indian hedgehog
26	1171	56.1	462	1 W48735	Human mutated soni
27	1167	55.9	437	1 R80071	Rat Vhh-1, Nucleic
28	1158	55.3	425	1 W61487	Chicken Sonic hedgehog
29	1154	55.3	416	1 R97658	Zebrafish tlggy-wi
30	1154	55.3	416	1 W61485	Zebrafish tlggy-wi
31	1154	55.3	416	1 W94475	Zebrafish tlggy-wi
32	1154	55.3	416	1 W97771	Zebrafish tlggy-wi
33	1154	55.3	416	1 Y05518	Zebrafish tlggy-wi
34	1126	53.9	411	1 W94470	Mouse Ih hedgehog

35	1126	53.9	411	1 W97767	Mouse Indian hedgehog
36	1126	53.9	411	1 Y05512	Mouse Indian hedgehog
37	1120	53.6	418	1 W61486	Zebrafish sonic he
38	1069	51.2	418	1 R77340	Zebrafish sonic he
39	1069	51.2	418	1 W94472	Zebrafish Sh hedgehog
40	1069	51.2	418	1 W97769	Zebrafish Sonic he
41	1069	51.2	418	1 Y05514	Zebrafish Sonic he
42	1045	50.0	198	1 W79597	Human Desert hedgehog
43	989	47.4	191	1 W79598	Human Desert hedgehog
44	964	46.2	182	1 W79596	Human Desert hedgehog
45	956.5	45.8	336	1 R77343	Mouse Indian hedgehog

ALIGNMENTS

RESULT 1
ID W97764
AC W97764; standard; Protein; 396 AA.
DT 21-MAY-1999 (first entry)
DE Human Desert hedgehog (Dhh) protein.
KW Sesert hedgehog; Dhh protein; human; dopaminergic; GABA-ergic; ptc therapeutic; patched; signal transduction; Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW cerebral ischaemia; hypoxia; neuroprotective; therapy.
OS Homo sapiens.
PN W09904775-A2.
PD 04-FEB-1999.
PE 24-JUL-1998; U15419.
PF 24-JUL-1997; US-900220.
PI (ONTO-) ONTOGENY INC.
PI Mahanthappa NK, Miao N, Pang K, Wang M;
DR WPI: 99-142578/12.
NR N-PSDB: X07270.
PT Increasing the survival of neuronal, dopaminergic and GABA-ergic
PT cells - by using a ptc therapeutic such as a protein kinase
PT inhibitor, or an agent derived from hedgehog polypeptides, useful in
PT the treatment of Parkinson's disease
PS Claim 30: Page 97-99; 138pp; English.
CC This polypeptide is human Dhh Desert hedgehog protein. The
CC invention is based on the finding that hedgehog proteins are useful
CC as protective agents in the treatment and prophylaxis of
CC neurodegenerative disorders resulting from the loss of dopaminergic
CC and/or GABA-ergic neurons, or the general loss of tissue from the
CC substantia nigra. Exemplary disorders include Parkinson's disease,
CC Huntington's disease (both claimed), amyotrophic lateral sclerosis
CC and cerebral ischaemia. The invention relates to hedgehog
CC therapeutics (i.e. hedgehog polypeptides and gene therapy constructs
CC e.g. constructs encoding recombinant hedgehog polypeptides and
CC trans-activation constructs for altering hedgehog gene regulatory
CC sequences) and ptc therapeutics (i.e. agents which mimic the effect
CC of naturally occurring hedgehog proteins on patched signalling)
CC that are effective in both human and animal subjects. A bioactive
CC polypeptide comprising amino acids 23-198 of human Dhh is preferred.
CC The products can also be used for the maintenance of differentiated
CC neurons in cultures, and to enhance the implantation of such
CC neuronal cells in an animal. They can be used to prevent or treat
CC neurodegenerative conditions arising from the use of certain drugs,
CC and in the prevention and/or treatment of hypoxia, e.g. as a
CC neuroprotective agent.
SQ Sequence 396 AA;

Query Match 100.0%; Score 2088; DB 1; Length 396;
Best Local Similarity 100.0%; Pred. No. 1.1e-216; Indels 0; Gaps 0;
Matches 396; Conservative 0; Mismatches 0;

OY 1 MALLNLPLPLCLALLALPAOSCGRPVGRARRAROLVLLKQVPGVPTLGAS 60
DB 1 MALLNLPLPLCLALLALPAOSCGRPVGRARRAROLVLLKQVPGVPTLGAS 60
OY 61 GPAEGRVARGSERFDLVPNVNPDIIFKDENSGADRLMTERCKERVNALIAVNMMPG 120

```

DB 61 GAEGRVARSERFRDLVFNYPNDIIFKDENSGADRLMERCKERNALAIAYNMMPG 120
QY 121 VRLRVTEGMEDEGHHADSLHYEGRALDITTSDRDNKYGGLARLAVEAGFDMVYYESRN 180
DB 121 VRLRVTEGMEDEGHHADSLHYEGRALDITTSDRDNKYGGLARLAVEAGFDMVYYESRN 180
QY 181 HHVSVKADNSLAVRAGGCGPGNATVRLMSGERRKGLRELHRDWMVLAADASGRVPTPV 240
DB 181 HHVSVKADNSLAVRAGGCGPGNATVRLMSGERRKGLRELHRDWMVLAADASGRVPTPV 240
QY 241 LFLDRDLQRRASFVAVETEMPPRKLLTPMHLVFAARGPAPGDFAPYFARRLRAGDSV 300
DB 241 LFLDRDLQRRASFVAVETEMPPRKLLTPMHLVFAARGPAPGDFAPYFARRLRAGDSV 300
QY 301 LAPGDLRPARVARYAREAVGVFAPLTAHGTLVNDVLAACVAVLSESHQAHRAFAPL 360
DB 301 LAPGDLRPARVARYAREAVGVFAPLTAHGTLVNDVLAACVAVLSESHQAHRAFAPL 360
QY 361 RLHLALGALLPGGAVOPTGMHWSRLLYRLAEELG 396
DB 361 RLHLALGALLPGGAVOPTGMHWSRLLYRLAEELG 396

RESULT 2
ID X05517 standard; Protein: 396 AA.
AC X05517;
DE 05-JUL-1999 (first entry)
DE Human Desert hedgehog protein Dhh.
KW Desert hedgehog; Dhh protein; human; hedgehog therapeutic;
KW ptc therapeutic; patched; signal transduction; muscle atrophy;
KW cachaexia; muscular myopathy; myoblastic sarcoma; therapy.
OS Homo sapiens.
PN W09910004-A2.
PD 24-MAR-1999.
PF 28-AUG-1998; U17922.
PR 29-AUG-1997; US-057394.
PA 31adgen CS, Currie PD, Hughes SM, Ingham PW;
PI WPI: 99-243557/20.
DR N-PSDB: X25105.

PT A new method to regulate muscle growth
PS Disclosure; page 121-121; 130pp; English.
CC The present sequence is human Desert hedgehog protein Dhh. The
CC invention relates to a method for modulating the formation and/or
CC maintenance of muscle tissue by ecotopically contacting muscle
CC cells, especially muscle stem/progenitor cells, in vitro or in
CC vivo, with a hedgehog therapeutic (i.e. hedgehog polypeptides and
CC gene therapy constructs) or ptc therapeutic (i.e. a small organic
CC molecule that mimics the effect of hedgehog proteins on patched
CC signaling, or activates or potentiates patched signaling) in an
CC amount effective to alter the growth state of the treated cells.
CC Also claimed is a method for treatment or prevention of disorders
CC of, or surgical or cosmetic repair of, such muscle tissues, by
CC administering a hedgehog polypeptide or ptc therapeutic. The
CC disorder may be muscle atrophy, in particular skeletal muscle
CC atrophy or cardiac muscle atrophy, cachaexia, or muscular myopathy
CC (all claimed). The hedgehog polypeptide or ptc therapeutic can
CC inhibit growth of myoblastic-derived tissue to provide treatment of
CC hyperplastic or neoplastic tissue such as in
CC myoblastic sarcoma (also claimed). The hedgehog therapeutic
CC preferably comprises at least a bioactive extracellular portion of
CC a hedgehog protein (see X05510-19) encoded by a vertebrate hedgehog
CC gene (see X25098-107), especially a human hedgehog gene.
SO Sequence 396 AA;

```

```

Query Match 100.0%; Score 2088; DB 1; Length 396;
Best Local Similarity 100.0%; Pred. No. 1.1e-216;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALLTNLPCLCLALLALPAQSCGPGRGVGRRRYARKOLVPLLYKQFVPGVPERTLGAS 60

```

```

DB 61 GAEGRVARSERFRDLVFNYPNDIIFKDENSGADRLMERCKERNALAIAYNMMPG 120
QY 61 GAEGRVARSERFRDLVFNYPNDIIFKDENSGADRLMERCKERNALAIAYNMMPG 120
DB 61 GAEGRVARSERFRDLVFNYPNDIIFKDENSGADRLMERCKERNALAIAYNMMPG 120
QY 121 VRLRVTEGMEDEGHHADSLHYEGRALDITTSDRDNKYGGLARLAVEAGFDMVYYESRN 180
DB 121 VRLRVTEGMEDEGHHADSLHYEGRALDITTSDRDNKYGGLARLAVEAGFDMVYYESRN 180
QY 181 HHVSVKADNSLAVRAGGCGPGNATVRLMSGERRKGLRELHRDWMVLAADASGRVPTPV 240
DB 181 HHVSVKADNSLAVRAGGCGPGNATVRLMSGERRKGLRELHRDWMVLAADASGRVPTPV 240
QY 241 LFLDRDLQRRASFVAVETEMPPRKLLTPMHLVFAARGPAPGDFAPYFARRLRAGDSV 300
DB 241 LFLDRDLQRRASFVAVETEMPPRKLLTPMHLVFAARGPAPGDFAPYFARRLRAGDSV 300
QY 301 LAPGDLRPARVARYAREAVGVFAPLTAHGTLVNDVLAACVAVLSESHQAHRAFAPL 360
DB 301 LAPGDLRPARVARYAREAVGVFAPLTAHGTLVNDVLAACVAVLSESHQAHRAFAPL 360
QY 361 RLHLALGALLPGGAVOPTGMHWSRLLYRLAEELG 396
DB 361 RLHLALGALLPGGAVOPTGMHWSRLLYRLAEELG 396

RESULT 3
ID W79595 standard; Protein: 396 AA.
AC W79595;
DE 02-FEB-1999 (first entry)
DE Human Desert hedgehog protein precursor.
KW Desert hedgehog; Huhhh; human; monoclonal antibody.
OS Homo sapiens.
PN Key
PD Key Peptide Location/Qualifiers
FT 1..22 /label= Sig_peptide
FT 23..198 /label= Mat_protein
FT Protein
PN EP-874048-A2.
PD 28-OCT-1998.
PF 24-APR-1998; 303187.
PR 14-APR-1998; JP-117873.
PR 25-APR-1997; JP-121578.
PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
PI Ariyasu T, Nakamura S, Orita K;
DR WPI: 98-544642/47.
DR N-PSDB: V62395.

PT Human Desert hedgehog protein - and corresponding DNA and monoclonal
PT antibody
PS Claim 4; Page 24-25; 39pp; English.
CC This is the amino acid sequence of a precursor of a novel human
CC Desert hedgehog protein, as deduced from a cDNA clone (see V62395)
CC derived from cell line ARH-77 (ATCC CRL-1621). The mature form
CC (see W79593) of the hedgehog protein is also claimed, as are a
CC monoclonal antibody (Mab) that recognises the protein, a process
CC for producing the protein, and a method for detecting the protein.
CC The hedgehog protein is useful in establishment of hybridomas which
CC produce antibodies recognising the protein, and the Mab is useful
CC for detecting and purifying the protein. The hedgehog protein,
CC DNA and Mab can be used to elucidate hereditary morphological
CC abnormalities in humans to establish their treatments and diagnoses.
SO Sequence 396 AA;

```

```

Query Match 99.8%; Score 2083; DB 1; Length 396;
Best Local Similarity 99.5%; Pred. No. 3.7e-216;
Matches 394; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MALLTNLPCLCLALLALPAQSCGPGRGVGRRRYARKOLVPLLYKQFVPGVPERTLGAS 60

```

Dh 1 MALTLNLPICCLALLPAOSCGRGVGRRRYARKQVLPVLYKQFVPGVPERTIGAS 60
QY 61 GPAGRVARGSERPRLDVPNNPDIIFKDEENSGADRLMTERCKERNALAIAMNMPG 120
Dh 61 GPAGRVARGSERPRLDVPNNPDIIFKDEENSGADRLMTERCKERNALAIAMNMPG 120
QY 121 VRLNTEGMDGHHADSLHYEGRALDITTSDDRNKYGGLARLAVAGFDWYYESRN 180
Dh 121 VRLNTEGMDGHHADSLHYEGRALDITTSDDRNKYGGLARLAVAGFDWYYESRN 180
QY 181 HHVSVADNSLAVRAGCGPGNATVRLMSGERGLRELHGDVNLADAGRVVPTVL 240
Dh 181 HHVSVADNSLAVRAGCGPGNATVRLMSGERGLRELHGDVNLADAGRVVPTVL 240
QY 241 LFLDRDLORRASFAVETEMPRLKLLTPMHLVFAARGPAPGDFAPVFARRLRAGDSV 300
Dh 241 LFLDRDLORRASFAVETEMPRLKLLTPMHLVFAARGPAPGDFAPVFARRLRAGDSV 300
QY 301 LAPGDLARPARVARVAREEAVGFAPLTAHGTLLVNDVLAACYAVLESHOMHRAFAPL 360
Dh 301 LAPGDLARPARVARVAREEAVGFAPLTAHGTLLVNDVLAACYAVLESHOMHRAFAPL 360
QY 361 RLHLALGALPGAVOPTGMHWSRLYLRLAEELG 396
Dh 361 RLHLALGALPGAVOPTGMHWSRLYLRLAEELG 396

RESULT 4

R77345
ID R77345 standard; Protein; 396 AA.

AC R77345;
DE Mouse desert hedgehog protein.
KW Mouse; desert hedgehog protein; probe; primer; diagnostic;
KM nervous system disorder; gene therapy; antibody.
OS Mus musculus.
FH Key
FT peptide 1..22
FT /note= "signal peptide"
FT 23..28
FT /note= "conserved sequence (R77345)"
FN W0518856-A1.
PD 13-JUL-1995.
PE 30-DEC-1994; U14992.
PR 30-DEC-1993; US-176427.
PR 14-DEC-1994; US-356060.
PA (HARD) HARVARD COLLEGE.
PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
PI Ingham PW, McMahon AP, Tablin CJ;
DR N-PSDB: Q91642.
PT Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful
PT to treat degenerative nervous system disorder(s) and in gene
PT therapy.
PS Claim 17; Page 135-37; 210pp; English.
CC The sequence represents a mouse Indian hedgehog protein, homologous
CC to a Drosophila hedgehog protein (R77337), and is encoded by a cDNA
CC isolated by low stringency screening of a mouse genome DNA library.
CC Probes and primers derived from hedgehog sequences may be used as
CC diagnostic agents for neuromuscular, autonomic or central nervous
CC system disorders, and the gene may also be used in gene therapy.
CC Antibodies generated from the protein may be used as therapeutic or
CC research reagents.
SQ Sequence 396 AA;

Query Match 96.2%; Score 2008; DB 1; Length 396;
Best Local Similarity 96.5%; Pred. No. 4.5e-208;
Matches 382; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 MALTLNLPICCLALLPAOSCGRGVGRRRYARKQVLPVLYKQFVPGVPERTIGAS 60
Dh 1 MALTLNLPICCLALLPAOSCGRGVGRRRYARKQVLPVLYKQFVPGVPERTIGAS 60

QY 61 GPAGRVARGSERPRLDVPNNPDIIFKDEENSGADRLMTERCKERNALAIAMNMPG 120
Dh 61 GPAGRVARGSERPRLDVPNNPDIIFKDEENSGADRLMTERCKERNALAIAMNMPG 120
QY 121 VRLNTEGMDGHHADSLHYEGRALDITTSDDRNKYGGLARLAVAGFDWYYESRN 180
Dh 121 VRLNTEGMDGHHADSLHYEGRALDITTSDDRNKYGGLARLAVAGFDWYYESRN 180
QY 181 HHVSVADNSLAVRAGCGPGNATVRLMSGERGLRELHGDVNLADAGRVVPTVL 240
Dh 181 HHVSVADNSLAVRAGCGPGNATVRLMSGERGLRELHGDVNLADAGRVVPTVL 240
QY 241 LFLDRDLORRASFAVETEMPRLKLLTPMHLVFAARGPAPGDFAPVFARRLRAGDSV 300
Dh 241 LFLDRDLORRASFAVETEMPRLKLLTPMHLVFAARGPAPGDFAPVFARRLRAGDSV 300
QY 301 LAPGDLARPARVARVAREEAVGFAPLTAHGTLLVNDVLAACYAVLESHOMHRAFAPL 360
Dh 301 LAPGDLARPARVARVAREEAVGFAPLTAHGTLLVNDVLAACYAVLESHOMHRAFAPL 360
QY 361 RLHLALGALPGAVOPTGMHWSRLYLRLAEELG 396
Dh 361 RLHLALGALPGAVOPTGMHWSRLYLRLAEELG 396

RESULT 5

W94469
ID W94469 standard; Protein; 396 AA.

AC W94469;
DE Mouse Dhh hedgehog protein sequence.
KW Patched; hedgehog; ptc therapeutic; neuroprotective; neuronal cell;
KM brain infarction; cerebral infarction; transient ischaemic attack;
KM stroke; cerebral infarct volume; spinal cord; oedema; trauma;
KM haemorrhage; encephalomyelitis; coronary bypass; cerebral hypoxia.
OS Mus sp.
FN W09900117-A2.
PD 07-JAN-1999.
PE 26-JUN-1998; U13387.
PR 27-JUN-1997; US-883656.
PA (ONTO-) ONTOGENY INC.
PI Mahanahappa NK;
DR WPI: 99-095458/08.
DR N-PSDB: X16183.
PT Method for limiting damage to neurons caused by ischaemic or epoxic
PT conditions - is used for the treatment and prevention of e.g.
PT cerebral infarction, stroke and transient ischaemic attacks
PS Disclosure; page 65-66; 104pp; English.
CC A method has been developed for limiting the damage to neuronal cells by
CC ischaemic or epoxic conditions by administering a ptc (patched) -
CC therapeutic agent to reduce cerebral infarct volume (CIV). Damage to
CC neuronal cells can also be limited by administering a gene activation
CC construct which recombines with the genomic hedgehog gene to provide a
CC heterologous transcription regulator linked to the coding region of this
CC gene. Administration of the ptc therapeutic agent is used to protect
CC cerebral tissues against ischaemic injury; to treat cerebral infarct or
CC ischaemia, stroke (thrombotic or embolic) and transient ischaemic
CC attacks. It may also be used as a prophylactic in many other cases of
CC injury to the brain or spinal cord, oedema caused by trauma, haemorrhage
CC and encephalomyelitis, or in conjunction with (coronary bypass) surgery.
CC Treatment (which may be prophylactic) is used where ischaemic/epoxic
CC conditions may cause cerebral hypoxia, or progressive loss of neurons
CC due to oxygen depletion, including in patients with hypotension. The
CC treatment reduces CIV by at least 25, particularly at least 70.%. The
CC present sequence represents a hedgehog sequence given in the present
CC invention.
SQ Sequence 396 AA;

Query Match 96.2%; Score 2008; DB 1; Length 396;
Best Local Similarity 96.5%; Pred. No. 4.5e-208;
Matches 382; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 MALTNLPLCLALLALPAQSCGPGRGVRRRRARAKQVPLLYKQVPGVPERTLGAS 60
 DB 1 MALPASLPLCLALLALSAQSCGPGRGVRRRRYRKQVPLLYKQVPSMPERTLGAS 60
 QY 61 GAEGGVARGSERFRDLVNYNPDIIFKDEENSGADRLMTERCKERVNALAIVANNMWP 120
 DB 61 GAEGGVARGSERFRDLVNYNPDIIFKDEENSGADRLMTERCKERVNALAIVANNMWP 120
 QY 121 VALRTGEGWDEGHHADSLHYEGRALDITTSDRNRKYGGLARLAVAGFDWYYESRN 180
 DB 121 VALRTGEGWDEGHHADSLHYEGRALDITTSDRNRKYGGLARLAVAGFDWYYESRN 180
 QY 181 HHVSVKADNSLAVRAGCGCPGNATVRLMSGERKGLRELHGDWVLAADAGRVPPTVL 240
 DB 181 HHVSVKADNSLAVRAGCGCPGNATVRLMSGERKGLRELHGDWVLAADAGRVPPTVL 240
 QY 241 LFLDRLQRRASFVAETERPRLKLLTPWHLVFARGPAPAGDFAPFARRLRAGDSV 300
 DB 241 LFLDRLQRRASFVAETERPRLKLLTPWHLVFARGPAPAGDFAPFARRLRAGDSV 300
 QY 301 LAPGDLAPARVARAREAVGFAPLTAHGTLLVNDVLASCVAVLESHQNAHRAFP 360
 DB 301 LAPGDLAPARVARAREAVGFAPLTAHGTLLVNDVLASCVAVLESHQNAHRAFP 360
 QY 361 RLHALGALLPGAVOPTGMHYSRLYLRLAEELG 396
 DB 361 RLHALGALLPGAVOPTGMHYSRLYLRLAEELG 396

RESULT 6
 W97766 standard; Protein: 396 AA.
 AC W97766;
 DT 21-MAY-1999 (first entry)
 DE Mouse Desert hedgehog (Dhh) protein.
 KM Desert hedgehog; Dhh protein; mouse; dopaminergic; GABA-nergic;
 KM ptc therapeutic; patched; signal transduction; Parkinson's disease;
 KM Huntington's disease; amyotrophic lateral sclerosis;
 KM cerebral ischemia; hypoxia; neuroprotective; therapy.
 OS Mus sp.
 PN W09904775-A2.
 PD 04-FEB-1999.
 PF 24-JUL-1998; U15419.
 PR 24-JUL-1997; US-900220.
 PA (ONTO-) ONTOGENY INC.
 PI Mahantappa NK, Miao N, Pang K, Wang M;
 DR WPI: 99-142578/12.
 DR N-PSDB: X072272.
 PT Increasing the survival of neuronal, dopaminergic and GABA-nergic
 PT cells - by using a ptc therapeutic such as a protein kinase
 PT inhibitor, or an agent derived from hedgehog polypeptides, useful in
 PT the treatment of Parkinson's disease
 PS Disclosure: Page 85-86; 138pp; English.
 CC This polypeptide is mouse Desert hedgehog protein (Dhh). The
 CC invention is based on the finding that hedgehog proteins are useful
 CC as protective agents in the treatment and prophylaxis of
 CC neurodegenerative disorders resulting from the loss of dopaminergic
 CC and/or GABA-nergic neurons, or the general loss of tissue from the
 CC substantia nigra. Exemplary disorders include Parkinson's disease,
 CC Huntington's disease (both claimed), amyotrophic lateral sclerosis
 CC and cerebral ischemia. The invention relates to hedgehog
 CC therapeutics (i.e. hedgehog polypeptides and gene therapy constructs
 CC a.g. constructs encoding recombinant hedgehog polypeptides and
 CC trans-activation constructs for altering hedgehog gene regulatory
 CC sequences) and ptc therapeutics (i.e. agents which mimic the effect
 CC of naturally occurring hedgehog proteins on patched signaling)
 CC that are effective in both human and animal subjects. Human Ihh
 CC and Dhh polypeptides (see W97763-64) are preferred. The products
 CC can also be used for the maintenance of differentiated neurons in
 CC cultures, and to enhance the implantation of such neuronal cells in
 CC an animal. They can be used to prevent or treat neurodegenerative
 CC conditions arising from the use of certain drugs, and in the

CC prevention and/or treatment of hypoxia, e.g. as a neuroprotective
 CC agent.
 SQ Sequence 396 AA:
 Query Match 96.2%; Score 2008; DB 1; Length 396;
 Best Local Similarity 96.5%; Pred. No. 4.5e-208;
 Matches 382; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 MALTNLPLCLALLALPAQSCGPGRGVRRRRARAKQVPLLYKQVPGVPERTLGAS 60
 DB 1 MALPASLPLCLALLALSAQSCGPGRGVRRRRYRKQVPLLYKQVPSMPERTLGAS 60
 QY 61 GAEGGVARGSERFRDLVNYNPDIIFKDEENSGADRLMTERCKERVNALAIVANNMWP 120
 DB 61 GAEGGVARGSERFRDLVNYNPDIIFKDEENSGADRLMTERCKERVNALAIVANNMWP 120
 QY 121 VALRTGEGWDEGHHADSLHYEGRALDITTSDRNRKYGGLARLAVAGFDWYYESRN 180
 DB 121 VALRTGEGWDEGHHADSLHYEGRALDITTSDRNRKYGGLARLAVAGFDWYYESRN 180
 QY 181 HHVSVKADNSLAVRAGCGCPGNATVRLMSGERKGLRELHGDWVLAADAGRVPPTVL 240
 DB 181 HHVSVKADNSLAVRAGCGCPGNATVRLMSGERKGLRELHGDWVLAADAGRVPPTVL 240
 QY 241 LFLDRLQRRASFVAETERPRLKLLTPWHLVFARGPAPAGDFAPFARRLRAGDSV 300
 DB 241 LFLDRLQRRASFVAETERPRLKLLTPWHLVFARGPAPAGDFAPFARRLRAGDSV 300
 QY 301 LAPGDLAPARVARAREAVGFAPLTAHGTLLVNDVLASCVAVLESHQNAHRAFP 360
 DB 301 LAPGDLAPARVARAREAVGFAPLTAHGTLLVNDVLASCVAVLESHQNAHRAFP 360
 QY 361 RLHALGALLPGAVOPTGMHYSRLYLRLAEELG 396
 DB 361 RLHALGALLPGAVOPTGMHYSRLYLRLAEELG 396

RESULT 7
 Y05511 standard; Protein: 396 AA.
 AC Y05511;
 DT 05-JUL-1999 (first entry)
 DE Mouse Desert hedgehog protein Dhh.
 KM Desert hedgehog; Dhh protein; mouse; hedgehog therapeutic;
 KM ptc therapeutic; patched; signal transduction; muscle atrophy;
 KM cachexia; muscular myopathy; myoblastic sarcoma; therapy.
 OS Mus sp.
 PN W09910004-A2.
 PD 04-MAR-1999.
 PF 28-AUG-1998; U17922.
 PR 29-AUG-1997; US-057394.
 PA (ONTO-) ONTOGENY INC.
 PI Bladgen CS, Currie PD, Hughes SM, Ingham PW;
 DR WPI: 99-243557/20.
 DR N-PSDB: X25099.
 PT A new method to regulate muscle growth
 PT Disclosure: Page 112-113; 130pp; English.
 PS The present sequence is mouse Desert hedgehog protein Dhh. The
 CC invention relates to a method for modulating the formation and/or
 CC maintenance of muscle tissue by ectopically contacting muscle
 CC cells, especially muscle stem/progenitor cells, in vitro or in
 CC vivo, with a hedgehog therapeutic (i.e. hedgehog polypeptides and
 CC gene therapy constructs) or ptc therapeutic (i.e. a small organic
 CC molecule that mimics the effect of hedgehog proteins on patched
 CC signaling, or activates or potentiates patched signaling) in an
 CC amount effective to alter the growth state of the treated cells.
 CC Also claimed is a method for treatment or prevention of disorders
 CC of, or surgical or cosmetic repair of, such muscle tissues, by
 CC administering a hedgehog polypeptide or ptc therapeutic. The
 CC disorder may be muscle atrophy, in particular skeletal muscle
 CC atrophy or cardiac muscle atrophy, cachexia, or muscular myopathy
 CC (all claimed). The hedgehog polypeptide or ptc therapeutic can

CC may be used as diagnostic agents for neuromuscular, autonomic or
 CC central nervous system disorders, and the gene may also be used in
 CC gene therapy. Antibodies generated from the protein may be used
 CC as therapeutic or research reagents.
 SQ Sequence 437 AA;

Query Match 58.3%; Score 1217; DB 1; Length 437;
 Best Local Similarity 58.7%; Pred. No. 9,66-123;
 Matches 249; Conservative 53; Mismatches 82; Indels 40; Gaps 12;

QY 7 LPLCCLALLA-----LPAQSCGPGRGVGRRRYARKOLVPLLYKQFVGPVPTLGASG 61
 DB 4 LLARCFVLTLASSLLVCPGLACGPGRG-FGKRHRPK-LTPLAYKQFIPVAEKTIGASG 61
 QY 62 PAEGVARGSEFRDLVFNYPNDIIFKDENSGADRLMTERCKERNALATAVMMMPGV 121
 DB 62 REGKITRNSERFKELTPVNPDIIFKDEENTGADRLMTORCKDKLNALATSVMMQPGV 121
 QY 122 RLRTVEGMDGDHHDSDLYHSGRALDITTSDDRNKYGLLARLAVEAGFDMVYVESRHH 181
 DB 122 RLRTVEGMDGDHHDSDLYHSGRALDITTSDDRNKYGLLARLAVEAGFDMVYVESRHH 181
 QY 182 VHVSKADNSLAVRAGCGFPGNATVRLMSGERKGLRELHGRDWTYLAADASGRVPTPVLL 241
 DB 182 IHCVSKAENSVAAKSGCGFPGSATVHLEGGTKLYDLRPGDRVLAADOGRLTYSDFLT 241
 QY 242 FLDRDQRRASFVAVETMPPRKLLTPMHLVFAA---RGPARAGDFAPVFAARLRAG 297
 DB 242 FLDRBEGAKKVFYVETLEPRERLLTAAHLTFVAPHNDSCPTPGP---SALFASRVPG 298
 QY 298 DSVLA---PGGD-ALRPARVARVA-REEAVGFAPLTAHGTLLVNDVLAACYAVLESRHH 352
 DB 298 QRVYVAERGGDRRLPAVAHSVTLREEAGAYAPLTAHGTLLINRVLAACYAVLESRHH 358
 QY 353 AHRAFAPLRLHA-LGALLP-----GGAU-----GPT-GMHWYSLTLYLA 391
 DB 353 AHRAFAPLRLHA-LGALLP-----GGAU-----GPT-GMHWYSLTLYLA 391
 QY 392 EELL 395
 DB 419 TWLL 422

RESULT 10
 W94471 standard; Protein: 437 AA.
 ID W94471 standard; Protein: 437 AA.
 AC W94471;
 DT 29-APR-1999 (first entry)
 DE Mouse Shh hedgehog protein sequence.
 KW Patched; hedgehog; ptc therapeutic; neuroprotective; neuronal cell;
 KW brain infarction; cerebral infarction; transient ischemic attack;
 KW stroke; cerebral infarct volume; spinal cord; oedema; trauma;
 KW haemorrhage; encephalomyelitis; coronary bypass; cerebral hypoxia.
 OS Mus sp.
 PN W09900117-A2.
 PD 07-JAN-1999.
 PF 26-JUN-1998; 013387.
 PR 27-JUN-1997; US-883656.
 PA (ONTO-) ONTOGENY INC.
 PI Mahanahappa NK;
 DR WPI: 99-095458/08.
 N-PSDB: X16185.
 PT Method for limiting damage to neurons caused by ischemic or epoxic
 PT conditions - is used for the treatment and prevention of e.g.
 PT cerebral infarction, stroke and transient ischemic attacks
 PS Disclosure: Page 68-70: 104pp: English.
 CC A method has been developed for limiting the damage to neuronal cells by
 CC ischemic or epoxic conditions by administering a ptc (patched)
 CC therapeutic agent to reduce cerebral infarct volume (CIV). Damage to
 CC neuronal cells can also be limited by administering a gene activation
 CC construct which recombines with the genomic hedgehog gene to provide a
 CC heterologous transcription regulator linked to the coding region of this

CC gene. Administration of the ptc therapeutic agent is used to protect
 CC cerebral tissues against ischemic injury; to treat cerebral infarct or
 CC ischemia, stroke (thrombotic or embolic) and transient ischemic
 CC attacks. It may also be used as a prophylactic in many other cases of
 CC injury to the brain or spinal cord, oedema caused by trauma, hemorrhage
 CC and encephalomyelitis, or in conjunction with (coronary bypass) surgery.
 CC Treatment (which may be prophylactic) is used where ischemic/epoxic
 CC conditions may cause cerebral hypoxia, or progressive loss of neurons
 CC due to oxygen depletion, including in patients with hypotension. The
 CC treatment reduces CIV by at least 25, particularly at least 70%. The
 CC present sequence represents a hedgehog sequence given in the present
 CC invention.
 SQ Sequence 437 AA;

Query Match 58.3%; Score 1217; DB 1; Length 437;
 Best Local Similarity 58.7%; Pred. No. 9,66-123;
 Matches 249; Conservative 53; Mismatches 82; Indels 40; Gaps 12;

QY 7 LPLCCLALLA-----LPAQSCGPGRGVGRRRYARKOLVPLLYKQFVGPVPTLGASG 61
 DB 4 LLARCFVLTLASSLLVCPGLACGPGRG-FGKRHRPK-LTPLAYKQFIPVAEKTIGASG 61
 QY 62 PAEGVARGSEFRDLVFNYPNDIIFKDENSGADRLMTERCKERNALATAVMMMPGV 121
 DB 62 REGKITRNSERFKELTPVNPDIIFKDEENTGADRLMTORCKDKLNALATSVMMQPGV 121
 QY 122 RLRTVEGMDGDHHDSDLYHSGRALDITTSDDRNKYGLLARLAVEAGFDMVYVESRHH 181
 DB 122 RLRTVEGMDGDHHDSDLYHSGRALDITTSDDRNKYGLLARLAVEAGFDMVYVESRHH 181
 QY 182 VHVSKADNSLAVRAGCGFPGNATVRLMSGERKGLRELHGRDWTYLAADASGRVPTPVLL 241
 DB 182 IHCVSKAENSVAAKSGCGFPGSATVHLEGGTKLYDLRPGDRVLAADOGRLTYSDFLT 241
 QY 242 FLDRDQRRASFVAVETMPPRKLLTPMHLVFAA---RGPARAGDFAPVFAARLRAG 297
 DB 242 FLDRBEGAKKVFYVETLEPRERLLTAAHLTFVAPHNDSCPTPGP---SALFASRVPG 298
 QY 298 DSVLA---PGGD-ALRPARVARVA-REEAVGFAPLTAHGTLLVNDVLAACYAVLESRHH 352
 DB 298 QRVYVAERGGDRRLPAVAHSVTLREEAGAYAPLTAHGTLLINRVLAACYAVLESRHH 358
 QY 353 AHRAFAPLRLHA-LGALLP-----GGAU-----GPT-GMHWYSLTLYLA 391
 DB 353 AHRAFAPLRLHA-LGALLP-----GGAU-----GPT-GMHWYSLTLYLA 391
 QY 392 EELL 395
 DB 419 TWLL 422

RESULT 11
 W97768 standard; Protein: 437 AA.
 ID W97768 standard; Protein: 437 AA.
 AC W97768;
 DT 21-MAY-1999 (first entry)
 DE Mouse Sonic hedgehog (Shh) protein.
 KW Sonic hedgehog; Shh protein; mouse; dopaminergic; GABA-ergic;
 KW ptc therapeutic; patched; signal transduction; Parkinson's disease;
 KW Huntington's disease; amyotrophic lateral sclerosis;
 KW cerebral ischemia; hypoxia; neuroprotective; therapy.
 OS Mus sp.
 PN W099004775-A2.
 PD 04-FEB-1999.
 PF 24-JUL-1998; 015419.
 PR 24-JUL-1997; US-900220.
 PA (ONTO-) ONTOGENY INC.
 PI Mahanahappa NK, Miao N, Pang K, Wang M;
 DR WPI: 99-142578/12.
 N-PSDB: X07274.
 PT Increasing the survival of neuronal, dopaminergic and GABA-ergic
 PT cells - by using a ptc therapeutic such as a protein kinase

PT inhibitor, or an agent derived from hedgehog polypeptides, useful in
PT the treatment of Parkinson's disease
PS Disclosure; Page 89-91; 138pp; English.
CC This polypeptide is mouse Shh Sonic hedgehog protein. The
CC invention is based on the finding that hedgehog proteins are useful,
CC as protective agents in the treatment and prophylaxis of
CC neurodegenerative disorders resulting from the loss of dopaminergic
CC and/or GABA-ergic neurons, or the general loss of tissue from the
CC substantia nigra. Exemplary disorders include Parkinson's disease,
CC Huntington's disease (both claimed), amyotrophic lateral sclerosis
CC and cerebral ischaemia. The invention relates to hedgehog
CC therapeutics (i.e. hedgehog polypeptides and gene therapy constructs
CC e.g. constructs encoding recombinant hedgehog polypeptides and
CC trans-activation constructs for altering hedgehog gene regulatory
CC sequences) and ptc therapeutics (i.e. agents which mimic the effect
CC of naturally occurring hedgehog proteins on patched signalling)
CC that are effective in both human and animal subjects. Human inh
CC and dh polypeptides (see W97763-64) are preferred. The products
CC can also be used for the maintenance of differentiated neurons in
CC cultures, and to enhance the implantation of such neuronal cells in
CC an animal. They can be used to prevent or treat neurodegenerative
CC conditions arising from the use of certain drugs, and in the
CC prevention and/or treatment of hypoxia, e.g. as a neuroprotective
CC agent.
SQ Sequence 437 AA;

Query Match 58.3%; Score 1217; DB 1; Length 437;
Best Local Similarity 58.7%; Pred. No. 9.6e-123;
Matches 249; Conservative 53; Mismatches 82; Indels 40; Gaps 12;

QY 7 LLLPCLALLA-----LPAQSCGPGRGVGRRRYARKOLVPLLYKQFVPGVPTIGASG 61
DB 4 LLARCFVLTLASSLLVCGCLACGPGRG-FGKRHRPKK-LTPLAYKQFIPNVAEKTIGASG 61

QY 62 PAEGVAVGSGSRFRDLYNPNYNDIIFKDENSGADRLMTERCKEKNALATAVMNMPGV 121
DB 62 RYEGKITNSERFKLTNPYNDIIFKDENGADRLMTERCKEKNALATAVMNMPGV 121

QY 122 RLRYTEGDEGHHADSLHREGRALDITTSRDNRKTYGLLARLAVEAGFDMVYESRNH 181
DB 122 RLRYTEGDEGHHSEELHREGRAVDITTSRDNRKTYGLLARLAVEAGFDMVYESRNH 181

QY 182 VHSVYKADNSLAVRAGGCGFNATVRLMSGKRGRLRELRHGDVLAADASGRVPTPYLL 241
DB 182 IHCYSKAEINSVAAGSGCGFPGSATVHLEGGTKLVKDLRPDRVLAADQGRLLYSDFLT 241

QY 242 FLDRDLQRRASVAVETEMPRKLLTPMHLVFAA-----RGAPAPGDFAVFARLRAG 297
DB 242 FLDRDEGAKKVFYVETLEPRERLLTLAHLLEFAPHNDSEPTGP---SALFASRVPRG 298

QY 298 DSVLA---PGGD-ALRPARVARVA-REBAVGFAPLTAHGTLLVNDVLAACYAVLESHQW 352
DB 298 QRVYVAERGGDRLLPAAVHSVTLREBEAGAYAPLTAHGTLLIRVLAACYAVLESHW 358

QY 353 AHRAPAPRLRLHA-LGALLP-----GGAV-----QPT-GMHWYSLLYRLA 391
DB 353 AHRAPAPRLRLHA-LGALLP-----GGAV-----QPT-GMHWYSLLYRLA 391

QY 392 EELL 395
DB 419 TWLL 422

RESULT 12
Y05513
ID Y05513 standard; Protein; 437 AA.
AC Y05513;
DT 05-JUL-1999 (first entry)
DE Mouse Sonic hedgehog protein Shh.
KW Sonic hedgehog; Shh protein; mouse; hedgehog therapeutic;
ptc therapeutic; patched; signal transduction; muscle atrophy;
cachexia; muscular myopathy; myoblastic sarcoma; therapy.

OS Mus SP.
PN WO991004-A2.
PD 04-MAR-1999.
PF 28-AUG-1998; U17922.
PR 29-AUG-1997; US-057394.
PA (ONTO-) ONTOGENY INC.
PI Bladgen CS, Currie PD, Hughes SM, Ingham PW;
DR WPI; 99-243557/20.
DR N-PSDB; X25101.
PT A new method to regulate muscle growth.
PS Disclosure; Page 115-116; 130pp; English.
CC The present sequence is mouse Sonic hedgehog protein Shh. The
CC invention relates to a method for modulating the formation and/or
CC maintenance of muscle tissue by ectopically contacting muscle
CC cells, especially muscle stem/progenitor cells, in vitro or in
CC vivo, with a hedgehog therapeutic (i.e. hedgehog polypeptides and
CC gene therapy constructs) or ptc therapeutic (i.e. a small organic
CC molecule that mimics the effect of hedgehog proteins on patched
CC signalling, or activates or potentiates patched signalling) in an
CC amount effective to alter the growth state of the treated cells.
CC Also claimed is a method for treatment or prevention of disorders
CC of, or surgical or cosmetic repair of, such muscle tissues, by
CC administering a hedgehog polypeptide or ptc therapeutic. The
CC disorder may be muscle atrophy, in particular skeletal muscle
CC atrophy or cardiac muscle atrophy, cachexia, or muscular myopathy
CC (all claimed). The hedgehog polypeptide or ptc therapeutic can
CC inhibit growth of myoblastic-derived tissue to provide treatment of
CC hyperplastic or neoplastic growth of muscle tissue such as in
CC myoblastic sarcoma (also claimed). The hedgehog therapeutic
CC preferably comprises at least a bioactive extracellular portion of
CC a hedgehog protein (see Y05510-19) encoded by a vertebrate hedgehog
CC gene (see X25098-107), especially a human hedgehog gene.
SQ Sequence 437 AA;

Query Match 58.3%; Score 1217; DB 1; Length 437;
Best Local Similarity 58.7%; Pred. No. 9.6e-123;
Matches 249; Conservative 53; Mismatches 82; Indels 40; Gaps 12;

QY 7 LLLPCLALLA-----LPAQSCGPGRGVGRRRYARKOLVPLLYKQFVPGVPTIGASG 61
DB 4 LLARCFVLTLASSLLVCGCLACGPGRG-FGKRHRPKK-LTPLAYKQFIPNVAEKTIGASG 61

QY 62 PAEGVAVGSGSRFRDLYNPNYNDIIFKDENSGADRLMTERCKEKNALATAVMNMPGV 121
DB 62 RYEGKITNSERFKLTNPYNDIIFKDENGADRLMTERCKEKNALATAVMNMPGV 121

QY 122 RLRYTEGDEGHHADSLHREGRALDITTSRDNRKTYGLLARLAVEAGFDMVYESRNH 181
DB 122 RLRYTEGDEGHHSEELHREGRAVDITTSRDNRKTYGLLARLAVEAGFDMVYESRNH 181

QY 182 VHSVYKADNSLAVRAGGCGFNATVRLMSGKRGRLRELRHGDVLAADASGRVPTPYLL 241
DB 182 IHCYSKAEINSVAAGSGCGFPGSATVHLEGGTKLVKDLRPDRVLAADQGRLLYSDFLT 241

QY 242 FLDRDLQRRASVAVETEMPRKLLTPMHLVFAA-----RGAPAPGDFAVFARLRAG 297
DB 242 FLDRDEGAKKVFYVETLEPRERLLTLAHLLEFAPHNDSEPTGP---SALFASRVPRG 298

QY 298 DSVLA---PGGD-ALRPARVARVA-REBAVGFAPLTAHGTLLVNDVLAACYAVLESHQW 352
DB 298 QRVYVAERGGDRLLPAAVHSVTLREBEAGAYAPLTAHGTLLIRVLAACYAVLESHW 358

QY 353 AHRAPAPRLRLHA-LGALLP-----GGAV-----QPT-GMHWYSLLYRLA 391
DB 353 AHRAPAPRLRLHA-LGALLP-----GGAV-----QPT-GMHWYSLLYRLA 391

QY 392 EELL 395
DB 419 TWLL 422

RESULT 13

W61488
ID W61488 standard; Protein: 437 AA.
AC W61488;
DT 20-OCT-1998 (first entry)
DE Mouse sonic hedgehog (shh) protein.
KW Hedgehog polypeptide; neuronal cell proliferation; zebrafish. shh;
KW cholesterol biosynthesis; pituitary gland gene expression; mouse;
KW liggy-winkle hedgehog protein; twnh; sonic hedgehog; major limb trauma;
KW neuronal degeneration; nerve-sparing agent; Smith-Lemli-Opitz syndrome.
OS Mus sp.
PN MO9830576-A1.
PR 16-JUL-1998.
PR 07-OCT-1997; U15753.
PR 02-OCT-1997; US-061323.
PR 07-OCT-1996; US-729743.
PA (UYUO) UNIV JOHNS HOPKINS SCHOOL OF MEDICINE.
PI Beachy PA, Porter JA;
DR WPI: 98-399053/34.
PT New hedgehog-derived poly(peptide)s - used to develop products for
PT modulating proliferation or differentiation of neuronal cells,
PT cholesterol biosynthesis or transport or expression of pituitary
PT gland gene(s)
PS Claim 13: Pages 153-154; 210pp; English.
CC This represents a mouse sonic hedgehog (shh) protein sequence. The
CC invention provides methods and compounds for modulating proliferation or
CC differentiation of neuronal cells, cholesterol biosynthesis or transport
CC or expression of pituitary gland genes. The method for affecting
CC cholesterol biosynthesis or transport in a cell comprises contacting a
CC cell with a compound that affects hedgehog, thereby affecting cholesterol
CC biosynthesis or transport. The methods for inhibiting the neural inducing
CC activity of a hedgehog polypeptide in cells, and for inducing pituitary
CC gland gene expression utilizes sequences selected from a zebrafish
CC liggy-winkle hedgehog (twnh) protein (W61488), a zebrafish sonic hedgehog
CC (shh) protein (W61488), a chicken shh protein (W61487) or a mouse shh
CC protein (W61488). The products and methods provide for compounds which
CC can affect hedgehog activity. They can be used for treating disorders
CC which arise from neuronal degeneration or abnormal function. They can
CC also be used as nerve-sparing agents or in restoring or promoting
CC appropriate patterning during the healing of major limb trauma. They can
CC also be used for treating Smith-Lemli-Opitz syndrome. The products can
CC also be used for detection and diagnosis.
SO Sequence 437 AA:

Query Match 58.1%; Score 1214; DB 1; Length 437;
Best Local Similarity 58.5%; Pred. No. 2e-122;
Matches 248; Conservative 54; Mismatches 82; Indels 40; Gaps 12;
QY 7 LRLPCLALLA-----LPQSGGPGRGVGRRRYARKQVLPVLYKQFVGPVEBRTGASG 61
DB 4 LLAARFELVITLASSLVCPGLACGPGRG-FGKRHRPKK-LTPPLAYKQFIPVNAEKTGASG 61
QY 62 PAERGAARSEFRDLVPYNPDIIFKDEENSGADRLTECKERNAALAIYVMNMPGV 121
DB 62 RREGKTRISEFKELTPYNPDIIFKDEENSGADRLTECKERNAALAIYVMNMPGV 121
QY 122 RLRTVEGMEDEGHNAODSLHYEGRALDITTSDRNRKYGGLARLAVAGFDWYYESRNH 181
DB 122 KLRTEGMEDEGHNSSESLHYEGRAVDITTSDRNRKYGGLARLAVAGFDWYYESRNH 181
QY 182 VAVSYKADNSLAVRAGCGCPGNATVRLMSGERKGLRELHSGDMVYLAADSGRVRPVL 241
DB 182 IHCSYKAEKNSVAKSGCGCPGNATVRLMSGERKGLRELHSGDMVYLAADSGRVRPVL 241
QY 242 FLDRDQKARASFAVETEMPRLKLLTPWHLVFAA-----RGRAPARPGDPAFARLRMG 297
DB 242 FLDRDQKARASFAVETEMPRLKLLTPWHLVFAA-----RGRAPARPGDPAFARLRMG 297
QY 298 DSVLA---DGGD-ALRPAVARVA--REEAVGYEAPLTAHGTLLVNDVLAACYAVLESHOM 352
DB 298 QRYVYVAERGGDRLLPAVHSTLREEAGAVAPLTAHGTLLINNVLAACAVIEHSHOM 358
QY 353 AHRAFPRLRLAH-LGALLP-----GGAV-----OPT-GMHWYSRLLYRLA 391

DB 359 AHRAFPRLRLAH-LGALLP-----GGAV-----OPT-GMHWYSRLLYRLA 391
QY 392 EELL 395
DB 419 TWLL 422
RESULT 14
R77338
ID R77338 standard; Protein: 425 AA.
AC R77338;
DT 22-FEB-1996 (first entry)
DE Chicken sonic hedgehog protein.
KW Chicken; sonic hedgehog protein; stage 22/22 limb bud; probe;
KW primer; diagnostic; nervous system disorder; gene therapy;
KW antibody.
OS Gallus domesticus.
PN Key
PR Location/Qualifiers
PR Peptide 1..26
PR /note= "putative signal peptide"
FT peptide 27..32
FT /note= "conserved sequence (R77349)"
FT modified site 282
FT /note= "N-linked glycosylation site"
PN W09518856-A1.
PR 13-JUL-1995.
PR 30-DEC-1994; U14992.
PR 30-DEC-1993; US-176427.
PR 14-DEC-1994; US-356060.
PA (HARD) HARVARD COLLEGE.
PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
PI Ingham PW, McMahon AP, Tablin CJ;
DR N-PSDB: Q91636.
PT Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful
PT to treat degenerative nervous system disorder(s) and in gene
PT therapy.
PS Claim 17: Page 133-35; 210pp; English.
CC The sequence represents a chicken sonic hedgehog protein.
CC homologous to a Drosophila hedgehog protein (R77337), and is
CC encoded by a cDNA isolated from a stage 22/22 limb bud cDNA
CC library. Probes and primers derived from the sonic hedgehog gene
CC may be used as diagnostic agents for neuromuscular, autonomic or
CC central nervous system disorders, and the gene may also be used in
CC gene therapy. Antibodies generated from the protein may be used
CC as therapeutic or research reagents.
SO Sequence 425 AA:

Query Match 57.4%; Score 1198; DB 1; Length 425;
Best Local Similarity 59.4%; Pred. No. 1e-120;
Matches 244; Conservative 56; Mismatches 91; Indels 20; Gaps 10;
QY 1 MALLTNLP---LCLALLALPAQSGGPGRGVGRRRYARKQVLPVLYKQFVGPVEBRTL 57
DB 4 MLLTRILLVGFC--ALLVSSGLCGGPGRG-IGKRHRPKK-LTPPLAYKQFIPVNAEKT 59
QY 58 GASGPAERGAARSEFRDLVPYNPDIIFKDEENSGADRLTECKERNAALAIYVMN 117
DB 60 GASGPAERGAARSEFRDLVPYNPDIIFKDEENSGADRLTECKERNAALAIYVMN 119
QY 118 WPGVRLRTVEGMEDEGHNAODSLHYEGRALDITTSDRNRKYGGLARLAVAGFDWY 177
DB 120 WPGVRLRTVEGMEDEGHNSSESLHYEGRAVDITTSDRNRKYGGLARLAVAGFDWY 179
QY 178 SRNHVAVKADNSLAVRAGCGCPGNATVRLMSGERKGLRELHSGDMVYLAADSGR 237
DB 180 SKAHTHCSYKAEKNSVAKSGCGCPGNATVRLHSGDMVYLAADSGRDLRLYS 239
QY 238 PVLLFLDRDQKARASFAVETEMPRLKLLTPWHLVFAA--RGRAPARPGDPA--PVFARR 293
DB 240 DFLTFDRDQKARASFAVETEMPRLKLLTPWHLVFAA--RGRAPARPGDPA--PVFARR 299

QY 294 LRAGDS--VIAPGGDALRPARVARVA-REAVGVFAPLTAHGTLVNDVASCYAVIESH 350
 DB 300 VKPGORVYVGEQGQOLLPASVSHSVSLREASGAYAPLTAGQTLINRYLASCYAVIEEH 359
 QY 351 QMAHRAFAPLRLHA-LGALLPGGAV-----OPTGMHWSRLYLRLAEELL 395
 DB 360 SMAHMAFAFPRLAAGLLAALCPDGAIPTAATTTGTHWISRLYLRLGGSWL 410

RESULT 15

W94468
 ID W94468 standard; Protein; 425 AA.
 AC W94468;
 DT 29-APR-1999 (first entry)
 DE Chicken Shh hedgehog protein sequence.
 KM Patched; hedgehog; ptc therapeutic; neuroprotective; neuronal cell;
 KM brain infarction; cerebral infarction; transient ischemic attack;
 KM stroke; cerebral infarct volume; spinal cord; oedema; trauma;
 KM haemorrhage; encephalomyelitis; coronary bypass; cerebral hypoxia.
 OS Gallus sp.
 PN MO9900117-A2.
 PD 07-JAN-1999.
 PF 26-JUN-1998; U13387.
 PR 27-JUN-1997; US-883656.
 PA (ONTO-) ONTOGENY INC.
 PI Mahanthappa NK;
 DR WPI: 99-095458/08.
 DR N-PSDB: X16182.
 PT Method for limiting damage to neurons caused by ischemic or epoxic
 PT conditions - is used for the treatment and prevention of e.g.
 PT cerebral infarction, stroke and transient ischemic attacks
 PS Disclosure; Page 63-64; 104pp; English.
 CC A method has been developed for limiting the damage to neuronal cells by
 CC ischemic or epoxic conditions by administering a ptc (patched)
 CC therapeutic agent to reduce cerebral infarct volume (CIV). Damage to
 CC neuronal cells can also be limited by administering a gene activation
 CC construct which recombines with the genomic hedgehog gene to provide a
 CC heterologous transcription regulator linked to the coding region of this
 CC gene. Administration of the ptc therapeutic agent is used to protect
 CC cerebral tissues against ischemic injury; to treat cerebral infarct or
 CC ischaemia, stroke (thrombotic or embolic) and transient ischemic
 CC attacks. It may also be used as a prophylactic in many other cases of
 CC injury to the brain or spinal cord, oedema caused by trauma, haemorrhage
 CC and encephalomyelitis, or in conjunction with (coronary bypass) surgery.
 CC Treatment (which may be prophylactic) is used where ischemic/epoxic
 CC conditions may cause cerebral hypoxia, or progressive loss of neurons
 CC due to oxygen depletion, including in patients with hypotension. The
 CC treatment reduces CIV by at least 25, particularly at least 70%. The
 CC present sequence represents a hedgehog sequence given in the present
 CC invention.
 SO Sequence 425 AA;

Query Match 57.4%; Score 1198; DB 1; Length 425;

Best Local Similarity 59.4%; Pred. No. 1e-120;

Matches 244; Conservative 56; Mismatches 91; Indels 20; Gaps 10;

QY 1 MALLTNLTP---LCCIALALAPAOSSCGRGPRVGRARRARAKQVPLLYKQVPGVPERTL 57
 DB 4 MLLTLRLVGFIC--ALTVSSGLTCGPRG-IGRRHRPK-LTPPLAYKQFIPNVAEKTL 59
 QY 58 GASGPAEGRVARGSERFDLYPNTNPDIIFDEENSGADRLMTERCKERNVALAIIVNM 117
 DB 60 GASGRYBEKTRNSERFELTPNPNPDILIFKDEENTGADRLMTQCKDKLNALISVNMQ 119
 QY 118 WPGVRLRVTEGMDGHHADSLHYGRALDITTSDRDRNKYGLLARLAVEAGFDWVYE 177
 DB 120 WPGVRLRVTEGMDGHHSESLHYEGRADVITTSDRDRSKYGMRLAVEAGFDWVYE 179
 QY 178 SRNHVHSVKAADNSLAVAGCGFPGNATVRLMSGERKGLRELHBDWYLAADASGRVPT 237
 DB 180 SKAHHCYSKAVENSVAAKSGCGFPGSATVHLHGQTKLVKDLSPGDRVLAADADGRLLYS 239

QY 238 PVLLFDRLQIRASVFAVETEMPRKLLTPMHLVFAA--RGPARAPGDFA--PVFARR 293
 DB 240 DFLFLDRMDSSKRLFYVETTRQPARARLLTAHLLFVAPQHNSAATGSGQALFASN 299
 QY 294 LRAGDS--VIAPGGDALRPARVARVA-REAVGVFAPLTAHGTLVNDVASCYAVIESH 350
 DB 300 VKPGORVYVGEQGQOLLPASVSHSVSLREASGAYAPLTAGQTLINRYLASCYAVIEEH 359
 QY 351 QMAHRAFAPLRLHA-LGALLPGGAV-----OPTGMHWSRLYLRLAEELL 395
 DB 360 SMAHMAFAFPRLAAGLLAALCPDGAIPTAATTTGTHWISRLYLRLGGSWL 410

Search completed: June 5, 2000, 08:16:20
 Job time: 2585 sec

Thu Jun 8 15:52:54 2000

us-08-900-220-17.rag

Page 10

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2000, 08:16:26 ; Search time 26.41 Seconds

(without alignments)
1078.998 Million cell updates/sec

Title: US-08-900-220-16

Sequence: 1 MSPARLRPRHFCVLVLLIL.....GRLLEGGSFHPLGMSGAGS 411

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1539	70.8	406	13	057567 notophthalm
2	1285	59.1	434	13	057404 pleurodeles
3	1253.5	58.1	414	13	09w709 parallelthy
4	1086.5	50.0	442	13	073803 fugu rubrip
5	1076.5	49.5	415	5	017499 branchiosto
6	914	42.0	410	5	061676
7	864	41.0	199	11	09w729
8	701	32.2	139	6	09x516
9	681	31.3	138	13	09w6c1
10	665	30.6	150	13	09ygv7
11	597	27.4	185	5	096699
12	510.5	23.5	129	11	09wmp6
13	472	21.7	119	13	042128
14	397	18.3	88	13	09y6c3
15	389	17.9	80	13	042441
16	239	11.0	137	13	042234
17	174.5	8.0	615	5	091573
18	173.5	8.0	868	5	09xv14
19	170.5	7.8	1021	5	09xuv2
20	161	7.4	790	5	022872

21	161	7.4	1226	5	021835	021835 caenorhabd1
22	154.5	7.1	1207	5	021535	021535 caenorhabd1
23	148.5	6.8	481	5	045992	045992 caenorhabd1
24	148	6.8	485	5	094128	094128 caenorhabd1
25	140	6.4	557	5	094129	094129 caenorhabd1
26	138.5	6.4	550	5	094130	094130 caenorhabd1
27	138.5	6.4	629	5	045273	045273 caenorhabd1
28	123.5	5.7	205	5	023193	023193 caenorhabd1
29	105	4.8	1997	2	005647	005647 streptomyc
30	102	4.7	751	5	021754	021754 caenorhabd1
31	100.5	4.6	478	5	043955	043955 leishmania
32	99	4.6	1013	2	053499	053499 mycobacteri
33	98.5	4.5	316	2	0928V4	0928V4 streptomyc
34	98	4.5	1222	1	09y6K2	09y6K2 aeropyrum p
35	97	4.5	1551	4	014160	014160 homo sapien
36	97	4.5	7463	2	0924X6	0924X6 streptomyc
37	96.5	4.4	54	13	042233	042233 coturnix co
38	96.5	4.4	647	12	039278	039278 equine hept
39	96	4.4	8563	2	054297	054297 streptomyc
40	95.5	4.4	299	4	09y6K2	09y6K2 homo sapien
41	95.5	4.4	846	2	053152	053152 mycobacteri
42	95	4.4	950	2	050470	050470 mycobacteri
43	95	4.4	1616	2	096285	096285 mycobacteri
44	95	4.4	2117	2	092429	092429 streptomyc
45	94.5	4.3	257	4	09y204	09y204 homo sapien

ALIGNMENTS

RESULT 1
ID 057567 PRELIMINARY; PRT; 406 AA.
AC 057567;
DT 01-JUN-1998 (TEMBLrel. 06, Created)
DT 01-JUN-1998 (TEMBLrel. 06, Last sequence update)
DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)
DE HEDGEHOG SEGMENT POLARITY HOMOLOG.
OS Notophthalmus viridescens (Eastern newt) (Triturus viridescens).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC Batrachia; Caudata; Salamandroidea; Salamandridae; Notophthalmus.
RN [1]
RP SEQUENCE FROM N.A.
RA STARK D.R., GATES P.B., BROCKES J.P., FERRETTI P.;
RL Dev. Dyn. 0:0-0(1998).
DR EMBL: AF047466; AAC03108.1; -
DR HSSP: Q62226; 1VHH.
DR PFAM: PF01085; HH_signal; 1.
DR PFAM: PF01079; HntL; 1.
DR PRINTS: PR00632; SONICHOG.
SQ SEQUENCE 406 AA; 45072 MW; 9D0FFA76 CRC32;

Query Match 70.8%; Score 1539; DB 13; Length 406;
Best Local Similarity 74.1%; Pred. No. 2.2e-117;
Matches 292; Conservative 31; Mismatches 69; Indels 2; Gaps 2;
Db 15 LLLGGPGLGCGPGRVVGRRPRPR-LIPUSYKFLVHPFKTLGASGRYEGKTAARNSR 73
QY 18 LLLVVPAMGCGPGRVVGRRPRPRKTVLAYKQFSPNVPEKTTGASGRYEGKTAARNSR 77
Db 15 LLLGGPGLGCGPGRVVGRRPRPR-LIPUSYKFLVHPFKTLGASGRYEGKTAARNSR 73
QY 78 FKEILTVPNPDIIKDEENTGADRLMTQRCRDLNSLAISVMNQPVKRLVTEGMEDG 137
Db 74 FKEILTVPNPDIIKDEENTGADRLMTQRCRDLNSLAISVMNQPVKRLVTEGMEDG 133
QY 138 HHSSESLHYEGRAVDITTSRDRNKKYGLARLAVAGFDWYVESKAVHCSVKSSESA 197
Db 134 HHSSESLHYEGRAVDITTSRDRNKKYGLARLAVAGFDWYVESKAVHCSVKSSESA 193
QY 198 AKTGCGPPARALATLESGEKIPIADLEPGHRLVCMDEGGRYTSDFLFDROSTAVKEF 257
Db 194 AKTGCGPPARALATLESGEKIPIADLEPGHRLVCMDEGGRYTSDFLFDROSTAVKEF 253


```

QY 258 QVETODPPRRRLATPAHLFTADNHTEPARFRATFASHVOPGOYVYAGVPLOPARV 317
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 254 YVETDPPRRRLATPAHLFTADNHTEPARFRATFASHVOPGOYVYAGVPLOPARV 313
QY 318 AAVSTHVALGAVAPLTKHGTLVEDVAVSCFAAVADHHLAQLAFWPLRLFHSLANGSWTP 377
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 314 VSTTGTDSGAVAPLTKHGTLVEDVAVSCFAAVADHHLAQLAFWPLRLFHSHGRETPR 373
QY 378 GEGVHWYPOLLYRLGRLLLEEGSFHPLGMSGAGS 411
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 374 -EGMHWYSLLYRLGRLLLEEGSFHPLGMSGAGS 406

```

```

RESULT 2
ID 057404 PRELIMINARY: PRT: 434 AA.
AC 057404;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE SONIC HEDGEHOG-RELATED PROTEIN.
GN PW-SHH.
OS Pleurodeles waltlilii (Iberian ribbed newt).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC Batrachia; Caudata; Salamandroides; Salamandridae; Pleurodeles.
RN [1]
RP SEQUENCE FROM N.A.
RA CAUDIT X., NICOLAS S., LE PARCO Y.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF005332; AAB94412.1;
DR HSSP: 062236; 1VHH.
DR PFAM: PF01085; HH_signal; 1.
DR PFAM: PF01079; Hnt; 1.
DR PRINTS: PRO0632; SONICHHOG.
SQ SEQUENCE 434 AA; 48421 MW; AFFEDDF3 CRC32;

```

Query Match 59.1%; Score 1285; DB 13; Length 434;

Best Local Similarity 61.2%; Pred. No. 1.1e-96;

Matches 255; Conservative 46; Mismatches 90; Indels 26; Gaps 5;

```

QY 20 LVYPAAMGCGPGRVGSRPRPKLVPLAYKQSPNVPEKTGASRGYEGKIARSSRRK 79
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 19 LVYPLGCGPGRGICQRRR-PQKLTPLAYKQIPNVSEKTLGASRGYEGKIARSSRRK 77
QY 80 ELTPNPNPDIIFKDEENTGADRLMTQRCRDLNSLAISVNMOPGVKRLVTEGMDGDGH 139
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 78 ELTPNPNPDIIFKDEENTGADRLMTQRCRDLNSLAISVNMOPGVKRLVTEGMDGDGH 137
QY 140 SEESLHYEGRAVDITTSDBRNKYGILLARLAVAGFDWYTESKAHVCSYKSEHSAAK 199
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 138 SEESLHYEGRAVDITTSDBRNKYGILLARLAVAGFDWYTESKAHVCSYKSEHSAAK 197
QY 200 TGSCFPAGAOVRLBSGARVALSAVRPGDRLVLAGEDGSPFSDVLIIFLDREPRRLAFOV 259
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 198 SGCGCFGSAVTIEQGVRIYKGLKGNRYLADVDEGRILYSPILFPMDEEETAKRVFY 257
QY 260 IETQDPPRRRLATPAHLFTADNHTEPAA-----RFRATFASHVOPGOYVYAG 308
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 258 IETSPRRRLATPAHLFTADNHTEPAA-----RFRATFASHVOPGOYVYAG 317
QY 309 VP--GIQPARVAVSTHVALGAVAPLTKHGTLVEDVAVSCFAAVADHHLAQLAFWPLRL 366
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 318 REGRGREATVDVYLEEAAGAVAPYAHGTVIDRYLSCYAVIEHSHVAHMAFAFLRV 377
QY 367 -----FSLANGSWTP-----GEGVHWYPOLLYRLGRLLLEEGSFHPLGMSGAGS 411
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 378 GEGVHWYPOLLYRLGRLLLEEGSFHPLGMSGAGS 434

```

```

RESULT 3
Q9W709 PRELIMINARY: PRT: 414 AA.
ID Q9W709

```

```

AC Q9W709;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE SONIC HEDGEHOG.
GN SHH.
OS Paralicthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
OC Pleuronectiformes; Pleuronectoides; Bothidae; Paralicthys.
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE: 99238226.
RA SUZUKI T., ICHIRO O., KUROKAWA T.;
RT "Retinoic acid given at late embryonic stage depresses sonic hedgehog
RT and Hoxd-4 expression in the pharyngeal area and induces skeletal
RT malformation in flounder (Paralicthys olivaceus) embryos."
RL Dev. Growth Differ. 41:143-152(1999).
DR EMBL: AB029748; BAA82360.1;
SQ SEQUENCE 414 AA; 45945 MW; ELFB12EE CRC32;

```

Query Match 58.1%; Score 1263.5; DB 13; Length 414;

Best Local Similarity 62.6%; Pred. No. 5.6e-95;

Matches 251; Conservative 47; Mismatches 92; Indels 11; Gaps 6;

```

QY 16 LLLLVPAAMGCGPGRVGSRPRPKLVPLAYKQSPNVPEKTGASRGYEGKIARSS 75
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 12 VLSLIVSSGCGGPGRGYG-RRRHPKLTPLAYKQIPNVSEKTLGASRGYEGKIARSS 70
QY 76 ERFKELTPYNTDIIFKDEENTGADRLMTQRCRDLNSLAISVNMOPGVKRLVTEGMD 135
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 71 ERFKELTPYNTDIIFKDEENTGADRLMTQRCRDLNSLAISVNMOPGVKRLVTEGMD 130
QY 136 DGHSEESIHYEGRAVDITTSDBRNKYGILLARLAVAGFDWYTESKAHVCSYKSEHS 195
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 131 DGHSEESIHYEGRAVDITTSDBRNKYGILLARLAVAGFDWYTESKAHVCSYKSEHS 190
QY 196 AAKTGGCFPAGAOVRLBSGARVALSAVRPGDRLVLAGEDGSPFSDVLIIFLDREPRRL 255
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 191 VAAKGGCGPGSSVTYLDQGTCKPKVALQTDGRVLAADHAGQPVYTDIFMDQOSTRR 250
QY 256 AQVETODPPRRRLATPAHLFTA-DNHTEPAR-FRATFASHVOPGOYVYAGVPLO 313
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 251 LEVYLET-DSGKILTPAHLFTVGHNSSTERRAHNGMSAVFSAQVPOQYFVLDIAERQ 309
QY 314 PARVAVSTHVALGAVAPLTKHGTLVEDVAVSCFAAVADHHLAQLAFWPLRLFH---SL 370
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 310 PVTAKRIYQHEHGSFAPYTAGTYVVDVYVLAACVAVIODEHLALAPVRLAHVWSL 369
QY 371 AWGSWTPG---EGVHWYPOLLYRLGRLLLEEGSFHPLGMSG 407
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 370 LFSQPOASAQKQGVHWYSKILYQLGTWLDSSHSHPLGMS 410

```

```

RESULT 4
ID 073803 PRELIMINARY: PRT: 442 AA.
AC 073803;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE FUGU HEDGEHOG.
GN FHR.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
OC Tetraodontiformes; Tetraodontidae; Tetraodontidae; Fugu.
RN [1]
RP SEQUENCE FROM N.A.
RA GELNER K., BRENNER S.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF056116; AAC34384.1;

```



```

QY      310 PELGAPAAVASTHVALNGVAFLTKHGTLVEDVDVASCPRAAADHLADLAPEPLRF--- 368
          | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db      308 RGVARKKAVSVTRTGHHTNVAAPTNRQGSIVLDGVAVSSYAMRMDEWVAHASFAPIMWTYY 367
          | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY      368 --HSIANGSWTPG--EGVHWMPOLLYRLGRLLLEEGSF 401
          | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      368 ISHMGIITDITDGEOEQRHYHWYTGGILKKIKKYMSDLKF 405
          | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT   7
O9WV29    PRELIMINARY;           PRT:   177 AA.
AC O9WV29.
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DI 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DR INDIAN HEDGEHOG PROTEIN (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
   [1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY.
RA GARCES P.L., MEYER R.A. JR., BROWN C.A., PRICE D.K.;
   "Indian hedgehog in rat.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF162914; AAd45372.1; -.
FT NON_TER 1
FT TER 1
SQ SEQUENCE 177 AA; 19739 MW; 24CFE1044 CRC32;
```

	Query Match	41.6%	Score 904	DB 11	length 177
	Best Local Similarity	96.6%	Pred. No. 2.9e-66		
	Matches 171	Conservative 4	Mismatches 2	Indels 0	Gaps 0
QY	102	LMTQCRDLRLSLAIVSNQMPGVLRVTEGDEDEGHSEESLHYEGRAVDITTSRDRN	161		
Db	1	LMTQCRDLRLSLAIVSNQMPGVLRVTEGDEDEGHSEESLHYEGRAVDITTSRDRN	60		
QY	162	KYGLLARLAVAGEFDMVYYESKAHVHCYSKSHSAAKTGCGFPAGAQVRLBSGARVALS	221		
Db	61	KYGLLARLAVAGEFDMVYYESKAHVHCYSKSHSAAKTGCGFPAGAQVRLBSGARVALS	120		
QY	222	AVRPEDRLVANGEOSPTFSDVLIILDRBPHRLARFQVLETODPPRRRLATPAHLLF	278		
Db	121	AVRPEDRLVANGEOSPTFSDVLIILDRBPHRLARFQVLETODPPRRRLATPAHLLF	177		

RESULT	8	
09XS16		
ID		
AC	09XS16	PRELIMINARY;
AC	09XS16	PRT; 139 AA.
DT	01-NOV-1999	(TtEMBLrel. 12, Created)
DT	01-NOV-1999	(TtEMBLrel. 12, Last sequence update)
DT	01-NOV-1999	(TtEMBLrel. 12, Last annotation update)
DE	SONIC HEDGEHOG (FRAGMENT).	
OS	Bos taurus (Bovine).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;	
OC	Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;	
OC	Bovinae; Bos.	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=TOOTH GERM;	
RA	KIYAMA E., IWAMOTO M., OHNORI T., KURIKU K., WU C., OKURA T.,	
RA	BASHIR M.M., TUCKER T., PACIFICI M.,	
RT	"Development of Stratum intermedium and its Role as a Sonic Hedgehog	
RT	Signaling Structure during Odontogenesis."	
RL	Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AF144100; AAD33926.1; -.	
FT	NON_TER 1	
FT	NON_TER 139 139	

Seq	Sequence	139 AA	15961 MW	32D3F025 CRC32
Query Match		32.2%	Score 701	DB 6
Best Local Similarity		92.8%	Pred. No. 6.7e-50	Length 139
Matches 129	Conservative	7	Mismatches 3	Indels 0
				Gaps 0
OY	48	AYKQSPNVPEKTLGASGRYEKGIARRSEREFELTPNPNPDIIFDEENTGADRLMTORC	107	
Db	1	AYKQIPNVAEKTGLASGRYEKGIIRNSENREFELTPNPNPDIIFDEENTGADRLMTORC	60	
OY	108	KDRRLSLAISVNWQPGYKGLRRTGCDMDGDGHSEESLHYEGRAVDTTSDRNRKTYGLA	167	
Db	61	KDKLNALAISSVDDQPGYKGLRRTGCDMDGDGHSEESLHYEGRAVDTTSDRNRKTYGLA	120	
OY	168	RLAVAGGFWMYTESKAHV	186	
Db	121	RLAVAGGFWMYTESKAHI	139	

RESULT	9		
ID	Q9W6C1	PRELIMINARY;	PRT; 138 AA.
AC	Q9W6C1;		
DT	01-NOV-1999 (TrEMBLrel. 12, Created)		
DT	01-NOV-1999 (TrEMBLrel. 12, Last sequence update)		
DT	01-NOV-1999 (TrEMBLrel. 12, Last annotation update)		
DE	SONIC HEDGEHOG PROTEIN (FRAGMENT).		
GN	SHH1.		
OS	Eleutherodactylus coqui.		
OC	Eumetazoa; Chordata; Craniata; Vertebrata; Amphibia;		
OC	Batrachia; Anura; Neobatrachia; Bufonidae; Leptodactylidae;		
OC	Eleutherodactylus.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	CARL T.F., RICHARDSON M.K., OLSSON L., SCHLOSSER G., KLYMKOWSKI M.W.,		
RA	HAKEN J.;		
RT	"Differences in vertebrate limb development revealed by studies of the		
RT	direct developing frog <i>E. coqui</i> ."		
RL	Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF113403; AAD23436.1; -.		
FT	NON_TER	1	1
FT	NON_TER	138	138
SQ	SEQUENCE	138 AA; 15751 MW; 203E8060 CRC32;	

Query Match		31.3%	Score 681:	DB 13:	Length 138:
Best Local Similarity		91.2%	Pred. NO.	2.8e-48;	
Matches 125;	Conservative	7;	Mismatches	5;	Indels 0; Gaps 0;
QY	54	PNVPEKTLGAGRGREGKIARSSERFEKLTTPVNPDIIFPKDEENTGADSLMTQRCRDLNS	113		
Db	2	PNAVEKNTLGASGRREKGKTRRSERFEKLTTPVNSDIIIFKDEENTAAADRLMTQRCKDKLNA	61		
QY	114	LAIISWMPGVKIRVTEGWEDDGHHSEESLHYGRAVDITTSDDRNKKYGLLARLAVEA	173		
Db	62	LAIISVMNIMPVKLRVTEGWEDDGHHSEESLHYGRAVDITTSDDRNKKYGMLARLAVEA	121		
QY	174	GEDWVYESKAHVHCSV	190		
Db	122	GEDWVYESKAHIHCVS	138		
RESULT	10				
Q9YGV7					
ID	Q9YGV7	PRELIMINARY;	PRT;	150 AA.	
AC	Q9YGV7;				
DT	01-MAY-1999	(TREMBLrel. 10, Created)			
DT	01-MAY-1999	(TREMBLrel. 10, Last sequence update)			
DT	01-NOV-1999	(TREMBLrel. 12, Last annotation update)			
DE	SONIC HEDGEHOG (FRAGMENT).				
OS	Ambystoma mexicanum (Axo1oc1).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;				

OC Batrachia; Caudata; Salamandroidea; Ambystomidae; Ambystoma.
RN [1]
RP SEQUENCE FROM N.A.
RA TOROK M.A., IZPILZDA-BELMONTÉ J.C., GARDINER D.M., BRYANT S.V.,
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF031480; AAD18128.1; -
DR HSSP: Q62226; 1VHH. 1
FT NON_TER 1 150
FT NON_TER 1 150
SQ SEQUENCE 150 AA; 16599 MW; 93563298 CRC32;

Query Match 30.6%; Score 665; DB 13; Length 150;
Best Local Similarity 82.6%; Pred. No. 6.3e-47;
Matches 123; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

QY 88 DIFKDEENTGADRLMORCKDLNSLAISVMNOMPGVKLVTEGMDGHHSESLHYE 147
DB 1 DIFKDEENTGADRLMORCKDLNSLAISVMNOMPGVKLVTEGMDGHHSESLHYE 60
QY 148 GRAVDITSDRNRKYGILARLAVEAGFDWVYSEKAVHCSVSHSAAKTGCEFPAG 207
DB 61 GRAVDITSDRNRKYGILARLAVEAGFDWVYSEKAVHCSVSHSAAKTGCEFPAG 120
QY 208 AOVPLESGARVALSAVPRGDRVLAMGEDG 236
DB 121 AKYTLHGVTFRPVKDLRPRGDRVLAMGOG 149

RESULT 11

096699 PRELIMINARY; PRT; 185 AA.

AC 096699;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE HEDGEHOG PROTEIN (FRAGMENT).
GN HH.
OS Junonia coenia (Peacock butterfly) (Precis coenia).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Papilionoidea; Nymphalidae; Nymphalinae; Junonia.
RN [1]
RP SEQUENCE FROM N.A.
RA KEYS D.N., LEWIS D.L., SELEGUE J.E., PEARSON B.J., GOODRICH L.V.,
RA JOHNSON R.L., GATES J., SCOTT M.P., CARROLL S.B.;
RT "Recruitment of a hedgehog regulatory circuit in butterfly eyespot
evolution."
RL Science 0:0-0(1999).
DR EMBL: AF117742; AAD08931.1; -
DR HSSP: Q62226; 1VHH. 1
FT NON_TER 1 185
FT NON_TER 1 185
SQ SEQUENCE 185 AA; 20745 MW; 96A09B5A CRC32;

Query Match 27.4%; Score 597; DB 5; Length 185;
Best Local Similarity 65.3%; Pred. No. 2.9e-41;
Matches 111; Conservative 30; Mismatches 25; Indels 4; Gaps 3;

QY 63 ASGREKGIASSSEFKLTPNYNDITFKDEENTGADRLMORCKDLNSLAISVMNOM 122
DB 3 ASGREKGIASSSEFKLTPNYNDITFKDEENTGADRLMORCKDLNSLAISVMNOM 62
QY 123 PGVKLVTEGMDGHHSESLHYEGRVAVDITSDRNRKYGILARLAVEAGFDWVYSE 182
DB 63 PGVKLVTEGMDGHHSESLHYEGRVAVDITSDRNRKYGILARLAVEAGFDWVYSE 122
QY 183 KAHVHCYSKSHSAAKTG-GCPAGVAVRLESGARVALSAVPRGDRVLA 231
DB 123 KSYHCYSKSHSAAKTG-GCPAGVAVRLESGARVALSAVPRGDRVLA 169

RESULT 12

09WUP6 PRELIMINARY; PRT; 129 AA.

AC 09WUP6;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE DESERT HEDGEHOG PROTEIN (FRAGMENT).
GN HH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-SPRAGUE-DAWLEY;
RA GARGES P.L., MEYER R.A., JR., BROWN C.A., PRICE D.K.;
RT "Desert hedgehog in the rat."
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF148226; AAD31927.1; -
FT NON_TER 1 129
FT NON_TER 1 129
SQ SEQUENCE 129 AA; 14578 MW; CB8B2D40 CRC32;

Query Match 23.5%; Score 510.5; DB 11; Length 129;
Best Local Similarity 72.8%; Pred. No. 1.9e-34;
Matches 91; Conservative 20; Mismatches 13; Indels 1; Gaps 1;

QY 28 CGPGR-VVSGRRRPPKLVPLAYKOFSPNVPEKTGASGREKGIASSSEFKLTPNYN 86
DB 5 CGPGRGVGRRRVRKQVLPILYKQFVSPMERLGSAGREGVYTGSEFRVLVNYN 64
QY 87 PDIFKDEENTGADRLMORCKDLNSLAISVMNOMPGVKLVTEGMDGHHSESLHY 146
DB 65 PDIFKDEENTGADRLMORCKDLNSLAISVMNOMPGVKLVTEGMDGHHSESLHY 124
QY 147 EGRAV 151
DB 125 EGRAL 129

RESULT 13

042128 PRELIMINARY; PRT; 119 AA.

ID 042128;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE SHH, PARTIAL CDS (FRAGMENT).
GN ME-SHH.
OS Oryzias latipes (Medaka fish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Atherinomorpha;
OC Cyprinodontiformes; Adrianichthyoidae; Adrianichthyidae; Oryziinae;
OC Oryzias.
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-BBRR; TISSUE-WHOLE EMBRYO;
RA ARAKI K.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB007129; BAA22368.1; -
DR HSSP: Q62226; 1VHH. 1
FT NON_TER 1 119
FT NON_TER 1 119
SQ SEQUENCE 119 AA; 13179 MW; 1CAE5021 CRC32;

Query Match 21.7%; Score 472; DB 13; Length 119;
Best Local Similarity 70.6%; Pred. No. 2.3e-31;
Matches 84; Conservative 17; Mismatches 18; Indels 0; Gaps 0;

QY 128 RYTEGMDGHHSESLHYEGRVAVDITSDRNRKYGILARLAVEAGFDWVYSEKAVH 187
DB 128 RYTEGMDGHHSESLHYEGRVAVDITSDRNRKYGILARLAVEAGFDWVYSEKAVH 187

DB 1 RTEGMDGDGHFEESLHYEGRAVDITTSRDCKSKYGLTSLRLAVEAGFDWVYSEKAH1H 60
QY 188 CSYKSHSAAKTGGCFPPAGQVRLSEGAVALSAPRGDRIVLAMEGDSPTSDVLIIF 246
DB 61 CSYKASSAANKSGCGFPGSSTVLTENGTPQVPYKLPQPGDRIVLAAQYDGNFVYTDPIMF 119

RESULT 14

QY 09YGU3 PRELIMINARY; PRT; 88 AA.
AC 09YGU3;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE DESERT HEDGEHOG PROTEIN (FRAGMENT).
GN DH.
OS Brachydonto rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinidae; Rasbora; Danio.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 99051425.
RA AMORES A., FORCE A., YAN Y.-L., WANG Y.-L., FRITZ A., PRINCE V.,
RA HO R., AMERICA C., LANGELAND J., WESTERFIELD M., EKKER M.,
RA POSTLETHWAIT J.,
RT "Zebrafish hox clusters and vertebrate genome evolution."
RL Science 282:1711-1714 (1998).
DR EMBL; AF071236; AAD15931.1;
DR HSSP; 062226; 1VHH.
FT NON_TER 1 88
FT NON_TER 1 88
SQ SEQUENCE 88 AA; 10069 MW; 2C17DE49 CRC32;

Query Match 18.3%; Score 397; DB 13; Length 88;
Best Local Similarity 83.0%; Pred. No. 1.9e-25;
Matches 73; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 105 QRCCKRLNSLAISVMNQPGVKLRVTEGMDGDGHSESLHYEGRAVDITTSRDCKSKY 164
DB 1 QRCCKDLTKLAIIVMNMQPGVRLVTEAMDEGDHPPGSLHYEGRAVDITTSRDCKSKY 60
QY 165 LLARLAVEAGFDWVYSEKAHVCYKS 192
DB 61 LLARLAVEAGFDWVYSEKAHVCYKS 88

RESULT 15

QY 042441 PRELIMINARY; PRT; 80 AA.
AC 042441;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE SONIC HEDGEHOG (FRAGMENT).
OS Oryzias latipes (Medaka fish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Atherinomorphi;
OC Cyprinodontiformes; Adrianichthyoidae; Adrianichthyidae; Oryziinae;
OC Oryzias.
RN [1]
RP SEQUENCE FROM N.A.
RA KOESTER R., STICK R., LOOSLI F., WITTBRODT J.,
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; 297019; CAB09695.1;
DR HSSP; 062226; 1VHH.
DR PFM; PF01085; HH_Signal; 1.
FT NON_TER 1 80
FT NON_TER 1 80
SQ SEQUENCE 80 AA; 9180 MW; 59BE42B8 CRC32;

Query Match 17.9%; Score 389; DB 13; Length 80;
Best Local Similarity 90.0%; Pred. No. 7.3e-25;
Matches 72; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 93 DEENTGADRLMTQRCCKRLNSLAISVMNQPGVKLRVTEGMDGDGHSESLHYEGRAVD 152
DB 1 DEENTGADRLMTQRCCKRLNSLAISVMNQPGVKLRVTEGMDGDGHFEFLHYEGRAVD 60

QY 153 ITTSRDCKSKYGLTSLRLAVE 172
DB 61 ITTSRDCKSKYGLTSLRLAVE 80

Search completed: June 5, 2000, 08:19:30
Job time: 184 sec

Thu Jun 8 15:52:41 2000

us-08-900-220-16.rspt

Page 7

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2000, 07:33:15 ; Search time 38.4 Seconds
(without alignments)
253.515 Million cell updates/sec

Title: US-08-900-220-16

Perfect score: 2175
Sequence: 1 MSPARLRRLHFCVLVLLLL.....GRLLLESGSFHPLGSGAGS 411

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188963

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: A_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2175	100.0	411	W94474	Human 1bh hedgehog
2	2175	100.0	411	W07462	Human Indian hedgehog
3	2175	100.0	411	W05516	Human 1bh hedgehog
4	2041	93.8	411	W94470	Mouse 1bh hedgehog
5	2041	93.8	411	W97767	Mouse Indian hedgehog
6	2041	93.8	411	Y05512	Mouse Indian hedgehog
7	1686	77.5	336	R77343	Mouse Indian hedgehog
8	1646	75.7	312	R77344	Human Indian hedgehog
9	1379.5	58.8	437	W61488	Mouse sonic hedgehog
10	1376.5	58.7	437	R77339	Mouse sonic hedgehog
11	1276.5	58.7	437	W94471	Mouse Shh hedgehog
12	1276.5	58.7	437	W97768	Mouse sonic hedgehog
13	1276.5	58.7	437	Y05513	Mouse sonic hedgehog
14	1274	58.6	418	W61486	Zebrfish sonic hedgehog
15	1274	58.4	475	R77341	Human sonic hedgehog
16	1271	58.4	475	W94473	Human Shh hedgehog
17	1271	58.4	475	W97770	Human sonic hedgehog
18	1271	58.4	475	Y05515	Human sonic hedgehog
19	1270.5	58.4	425	R77338	Chicken sonic hedgehog
20	1270.5	58.4	425	W94468	Chicken Shh hedgehog
21	1270.5	58.4	425	Y05510	Chicken sonic hedgehog
22	1270.5	58.4	425	Y05510	Chicken sonic hedgehog
23	1264.5	58.1	462	W48736	Human mutated sonic hedgehog
24	1262.5	58.0	462	W48735	Human mutated sonic hedgehog
25	1227	56.4	418	R77340	Zebrfish sonic hedgehog
26	1227	56.4	418	W94472	Zebrfish Shh hedgehog
27	1227	56.4	418	W97769	Zebrfish sonic hedgehog
28	1227	56.4	418	Y05514	Zebrfish sonic hedgehog
29	1226.5	56.4	416	R97658	Zebrfish sonic hedgehog
30	1226.5	56.4	416	W61485	Zebrfish sonic hedgehog
31	1226.5	56.4	416	W94475	Zebrfish sonic hedgehog
32	1226.5	56.4	416	W97771	Zebrfish sonic hedgehog
33	1226.5	56.4	416	Y05518	Zebrfish sonic hedgehog
34	1224.5	56.3	437	R80071	Rat Vhh-1, Nucleic

35	1190.5	54.7	425	1	W61487	Chicken sonic hedgehog
36	1175	54.0	336	1	R77345	Mouse desert hedgehog
37	1175	54.0	336	1	W94469	Mouse Dh hedgehog
38	1175	54.0	336	1	W97766	Mouse Dh hedgehog
39	1175	54.0	336	1	Y05511	Mouse Dh hedgehog
40	1171.5	53.9	336	1	W97764	Human Dh hedgehog
41	1171.5	53.9	336	1	Y05517	Human Dh hedgehog
42	1166.5	53.6	336	1	W79595	Human Dh hedgehog
43	1153	53.0	374	1	W79594	Human Dh hedgehog
44	893.5	41.1	471	1	R77337	Drosophila hedgehog
45	890.5	40.9	221	1	W94477	Shh polypeptide sequence

ALIGNMENTS

RESULT 1	W94474	Standard; Protein; 411 AA.
ID	W94474	Standard; Protein; 411 AA.
AC	W94474	
DT	29-APR-1999	(first entry)
DE	Human 1bh hedgehog protein sequence.	
KW	Patched; hedgehog; ptc therapeutic; neuroprotective; neuronal cell; brain infarction; cerebral infarction; transient ischaemic attack; stroke; cerebral infarct volume; spinal cord; oedema; trauma; haemorrhage; encephalomyelitis; coronary bypass; cerebral hypoxia.	
OS	Homo sapiens.	
PN	W0900117-A2.	
PD	07-JAN-1999.	
PF	26-JUN-1998: U13387.	
PR	27-JUN-1997: U5-883656.	
PA	(ONTO-) ONTOGENET INC.	
PI	Mahantappa NK;	
DR	N-PSDB: X16188.	
DR	WPI: 99-095458/08.	
PT	Method for limiting damage to neurons caused by ischaemic or epoxic conditions is used for the treatment and prevention of e.g. cerebral infarction, stroke, and transient ischaemic attacks	
PT	Disclosure; Page 74-76; 104pp; English.	
PS	A method has been developed for limiting the damage to neuronal cells by therapeutic agent to reduce cerebral infarct volume (CIV). Damage to neuronal cells can also be limited by administering a gene activation construct which recombines with the genomic hedgehog gene to provide a heterologous transcription regulator linked to the coding region of this gene. Administration of the ptc therapeutic agent is used to protect or cerebral tissues against ischaemic injury, to treat cerebral infarct or ischaemia, stroke (thrombotic or embolic) and transient ischaemic attacks. It may also be used as a prophylactic in many other cases of injury to the brain or spinal cord, oedema caused by trauma, haemorrhage and encephalomyelitis, or in conjunction with (coronary bypass) surgery. Treatment (which may be prophylactic) is used where ischaemic/epoxic conditions may cause cerebral hypoxia, or progressive loss of neurons due to oxygen depletion, including in patients with hypotension. The treatment reduces CIV by at least 25, particularly at least 70%. The present sequence represents a hedgehog sequence given in the present invention.	
CC	Sequence 411 AA;	

Query Match 100.0%; Score 2175; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.1e-213;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY	1	MSPARLRRLHFCVLVLLVPAWCGGGRVGRPRPKVPLAYKQSPNVEKT	60
DB	1	MSPARLRRLHFCVLVLLVPAWCGGGRVGRPRPKVPLAYKQSPNVEKT	60
OY	61	LGASGRYEGKIASRSERFKELTPYNDIIFKDEENGAADRLMTORCKDLNLSIAISVMN	120
DB	61	LGASGRYEGKIASRSERFKELTPYNDIIFKDEENGAADRLMTORCKDLNLSIAISVMN	120
OY	121	QWGVKLRVTEGWDEGHSSESLHTEGRAVDITTSRDRNKKIGLLARLAVEAGFDWYI	180

DB 121 QMPGKLTETGDEDESESLHTEGRAVDITTSDRNRKNTGILARLAVAGDWMY 180
QY 181 ESKAHVHCVSSEHSAATGCGCPAGAOVRLSEGARVALSAVRPGDRLVLMGEGSPTF 240
DB 181 ESKAHVHCVSSEHSAATGCGCPAGAOVRLSEGARVALSAVRPGDRLVLMGEGSPTF 240
QY 241 SVLLFLDREPRRLRAFOVIEETODPPRLALTPAHLFTADNHTPEPARFRATFASHVP 300
DB 241 SVLLFLDREPRRLRAFOVIEETODPPRLALTPAHLFTADNHTPEPARFRATFASHVP 300
QY 301 GOYVYVAGVPGLOPARVAVSTHVALGAYAPLTKHGTIVVEDVVASCPAAVADHHLAOLA 360
DB 301 GOYVYVAGVPGLOPARVAVSTHVALGAYAPLTKHGTIVVEDVVASCPAAVADHHLAOLA 360
QY 361 FMPRLRFLSHLAWGSMTPGEGVHWYPTQLYRGLRLLLEGSEFHPLMGSGAGS 411
DB 361 FMPRLRFLSHLAWGSMTPGEGVHWYPTQLYRGLRLLLEGSEFHPLMGSGAGS 411
RESULT 2
W97763 standard; Protein: 411 AA.
ID W97763
AC W97763:
DE 21-MAY-1999 (first entry)
DE Human Indian hedgehog (Ihh) protein.
KW Indian hedgehog; Ihh protein; human; dopaminergic; GABA-nergic;
KW ptc therapeutic; patched; signal transduction; Parkinson's disease;
KW Huntington's disease; amyotrophic lateral sclerosis;
KW cerebral ischemia; hypoxia; neuroprotective; therapy.
OS Homo sapiens.
PN W09904775-A2.
PD 04-FEB-1999.
PF 24-JUL-1998; U15419.
PR 24-JUL-1997; US-900220.
PA (ONTO-) ONTOGENY INC.
PI Mahantappa NK, Miao N, Pang K, Wang M;
DR WPI: 99-142578/12.
DR N-PSDB: X07269.
PT Increasing the survival of neuronal, dopaminergic and GABA-nergic
PT cells - by using a ptc therapeutic such as a protein kinase
PT inhibitor, or an agent derived from hedgehog polypeptides, useful in
PT the treatment of Parkinson's disease.
PS Claim 30; Page 95-97; 138pp; English.
CC This polypeptide is human Indian hedgehog protein (Ihh). The
CC invention is based on the finding that hedgehog proteins are useful
CC as protective agents in the treatment and prophylaxis of
CC neurodegenerative disorders resulting from the loss of dopaminergic
CC and/or GABA-nergic neurons, or the general loss of tissue from the
CC substantia nigra. Exemplary disorders include Parkinson's disease,
CC Huntington's disease (both claimed); amyotrophic lateral sclerosis
CC and cerebral ischemia. The invention relates to hedgehog
CC therapeutics (i.e. hedgehog polypeptides and gene therapy constructs
CC e.g. constructs encoding recombinant hedgehog polypeptides and
CC trans-activation constructs for altering hedgehog gene regulatory
CC sequences) and ptc therapeutics (i.e. agents which mimic the effect
CC of naturally occurring hedgehog proteins on patched signaling)
CC that are effective in both human and animal Ihh is preferred.
CC polypeptide comprising amino acids 28-202 of human Ihh is preferred.
CC The products can also be used for the maintenance of differentiated
CC neurons in cultures, and to enhance the implantation of such
CC neuronal cells in an animal. They can be used to prevent or treat
CC neurodegenerative conditions arising from the use of certain drugs,
CC and in the prevention and/or treatment of hypoxia, e.g. as a
CC neuroprotective agent.
SQ Sequence 411 AA;

Query Match 100.0%; Score 2175; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 1,1e-213;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPARLRPLRHLPCVLLLVPAAMGCGPGRVSGRRPRPKVLAIKQSPNPPEKT 60

DB 1 MSPARLRPLRHLPCVLLLVPAAMGCGPGRVSGRRPRPKVLAIKQSPNPPEKT 60
QY 61 LGASRYEGKILARSSERKEKLPYNNPDIIFKDEENTGADRLMTORCDRLNSLAISVN 120
DB 61 LGASRYEGKILARSSERKEKLPYNNPDIIFKDEENTGADRLMTORCDRLNSLAISVN 120
QY 121 QMPGKLTETGDEDEHSESLHTEGRAVDITTSDRNRKNTGILARLAVAGDWMY 180
DB 121 QMPGKLTETGDEDEHSESLHTEGRAVDITTSDRNRKNTGILARLAVAGDWMY 180
QY 181 ESKAHVHCVSSEHSAATGCGCPAGAOVRLSEGARVALSAVRPGDRLVLMGEGSPTF 240
DB 181 ESKAHVHCVSSEHSAATGCGCPAGAOVRLSEGARVALSAVRPGDRLVLMGEGSPTF 240
QY 241 SVLLFLDREPRRLRAFOVIEETODPPRLALTPAHLFTADNHTPEPARFRATFASHVP 300
DB 241 SVLLFLDREPRRLRAFOVIEETODPPRLALTPAHLFTADNHTPEPARFRATFASHVP 300
QY 301 GOYVYVAGVPGLOPARVAVSTHVALGAYAPLTKHGTIVVEDVVASCPAAVADHHLAOLA 360
DB 301 GOYVYVAGVPGLOPARVAVSTHVALGAYAPLTKHGTIVVEDVVASCPAAVADHHLAOLA 360
QY 361 FMPRLRFLSHLAWGSMTPGEGVHWYPTQLYRGLRLLLEGSEFHPLMGSGAGS 411
DB 361 FMPRLRFLSHLAWGSMTPGEGVHWYPTQLYRGLRLLLEGSEFHPLMGSGAGS 411
RESULT 3
Y05516 standard; Protein: 411 AA.
ID Y05516
AC Y05516:
DE 05-JUL-1999 (first entry)
DE Human Indian hedgehog protein Ihh.
KW Indian hedgehog; Ihh protein; human; hedgehog therapeutic;
KW ptc therapeutic; patched; signal transduction; muscle atrophy;
KW cachexia; muscular myopathy; myoblastic sarcoma; therapy.
OS Homo sapiens.
PN W09910004-A2.
PD 04-MAR-1999.
PF 28-AUG-1998; U17922.
PR 29-AUG-1997; US-057394.
PA (ONTO-) ONTOGENY INC.
PI Bladgen CS, Currie PD, Hughes SM, Ingham PW;
DR WPI: 99-243557/20.
DR N-PSDB: X25104.
PT A new method to regulate muscle growth
PT disclosure: Page 120-121; 130pp; English.
PS The present sequence is human Indian hedgehog protein Ihh. The
PS invention relates to a method for modulating the formation and/or
PS maintenance of muscle tissue by ecotopically contacting muscle
PS cells, especially muscle stem/progenitor cells, in vitro or in
PS vivo, with a hedgehog therapeutic (i.e. hedgehog polypeptides and
PS gene therapy constructs) or ptc therapeutic (i.e. a small organic
PS molecule that mimics the effect of hedgehog proteins on patched
PS signaling, or activates or potentiates patched signaling) in an
PS amount effective to alter the growth state of the treated cells.
PS Also claimed is a method for treatment or prevention of disorders
PS of, or surgical or cosmetic repair of, such muscle tissues, by
PS administering a hedgehog polypeptide or ptc therapeutic. The
PS disorder may be muscle atrophy, in particular skeletal muscle
PS atrophy or cardiac muscle atrophy, cachexia, or muscular myopathy
PS (all claimed). The hedgehog polypeptide or ptc therapeutic can
PS inhibit growth of myoblastic-derived tissue or provide treatment of
PS hyperplastic or neoplastic growth of muscle tissue such as in
PS myoblastic sarcoma (also claimed). The hedgehog therapeutic
PS preferably comprises at least a bioactive extracellular portion of
PS a hedgehog protein (see Y05510-19) encoded by a vertebrate hedgehog
PS gene (see X25098-107), especially a human hedgehog gene.
SQ Sequence 411 AA;

Query Match 100.0%; Score 2175; DB 1; Length 411;

Best Local Similarity 100.0%; Pred. No. 1.1e-213;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MSPARLPRHLFCVLVLLLVPAAMCGPGRVVGSRRRPRKLVPLAYKQFSPNVPEKT 60
    |||||
Dd 1 MSPARLPRHLFCVLVLLLVPAAMCGPGRVVGSRRRPRKLVPLAYKQFSPNVPEKT 60
    |||||
QY 61 LGASGRREGKTAARSERFKELTPTNPNPDIIFKDEENTGADRLMTQRCDRNLNSLAISVMN 120
    |||||
Dd 61 LGASGRREGKTAARSERFKELTPTNPNPDIIFKDEENTGADRLMTQRCDRNLNSLAISVMN 120
    |||||
QY 121 QMPGVKRLVTEGMDGDHSEESLHYEGRAVDITTSDBDRKKYGLRLARLAEAGFDWYY 180
    |||||
Dd 121 QMPGVKRLVTEGMDGDHSEESLHYEGRAVDITTSDBDRKKYGLRLARLAEAGFDWYY 180
    |||||
QY 181 ESKAHVHCYSKSEHSAAKTGCGCPAGAOVRLSEGARVALSAVRPGDVLAMGEDGSPTE 240
    |||||
Dd 181 ESKAHVHCYSKSEHSAAKTGCGCPAGAOVRLSEGARVALSAVRPGDVLAMGEDGSPTE 240
    |||||
QY 241 SDVLIIFLDREPHRLRAFOVLETODPPRRLLTPAHLLEFTADNHTEPARFRATASHVQ 300
    |||||
Dd 241 SDVLIIFLDREPHRLRAFOVLETODPPRRLLTPAHLLEFTADNHTEPARFRATASHVQ 300
    |||||
QY 301 GOYLVAGVPGLOPARAAVSTHVALGAYAPLTKHGLTVEDVYASCFPAADHHLAQLA 360
    |||||
Dd 301 GOYLVAGVPGLOPARAAVSTHVALGAYAPLTKHGLTVEDVYASCFPAADHHLAQLA 360
    |||||
QY 361 FWPRLRFLHSLAMGSWTPGEGVHWYPOLLYRLGRLLLEGGSFHPLMGSGAGS 411
    |||||
Dd 361 FWPRLRFLHSLAMGSWTPGEGVHWYPOLLYRLGRLLLEGGSFHPLMGSGAGS 411
    |||||
```

RESULT 4
W94470 standard; Protein; 411 AA.
AC W94470; 1999 (first entry)
DE Mouse Ihn hedgehog protein sequence.
KW Patched; hedgehog; ptc therapeutic; neuroprotective; neuronal cell;
KW brain infarction; cerebral infarction; transient ischemic attack;
KW stroke; cerebral infarct volume; spinal cord; oedema; trauma;
KW haemorrhage; encephalomyelitis; coronary bypass; cerebral hypoxia.
OS Mus sp.
PN WO9900117-A2.
PD 07-JAN-1999.
PE 26-JUN-1998; U13387.
PR 27-JUN-1997; US-883656.
PA (ONTO-) ONTOGENY INC.
PI Mahanthappa NK;
DR N-PSDB; X16184.
PT Method for limiting damage to neurons caused by ischemic or epoxic
PT conditions - is used for the treatment and prevention of e.g.
PT cerebral infarction, stroke and transient ischemic attacks
PS Disclosure; Page 66-68; 104pp; English.
CC A method has been developed for limiting the damage to neuronal cells by
CC ischemic or epoxic conditions by administering a ptc (patched)
CC therapeutic agent to reduce cerebral infarct volume (CIV). Damage to
CC neuronal cells can also be limited by administering a gene activation
CC construct which recombines with the genomic hedgehog gene to provide a
CC heterologous transcription regulator linked to the coding region of this
CC gene. Administration of the ptc therapeutic agent is used to protect
CC cerebral tissues against ischemic injury; to treat cerebral infarct or
CC ischaemia, stroke (thrombotic or embolic) and transient ischemic
CC attacks. It may also be used as a prophylactic in many other cases of
CC injury to the brain or spinal cord, oedema caused by trauma, haemorrhage
CC and encephalomyelitis, or in conjunction with (coronary bypass) surgery.
CC Treatment (which may be prophylactic) is used where ischemic/epoxic
CC conditions may cause cerebral hypoxia, or progressive loss of neurons
CC due to oxygen depletion, including in patients with hypotension. The
CC treatment reduces CIV by at least 25, particularly at least 70%. The
CC present sequence represents a hedgehog sequence given in the present
CC invention.

SO Sequence 411 AA;

Query Match 93.8%; Score 2041; DB 1; Length 411;
Best Local Similarity 94.6%; Pred. No. 5.3e-200;
Matches 389; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

```
QY 1 MSPARLPRHLFCVLVLLLVPAAMCGPGRVVGSRRRPRKLVPLAYKQFSPNVPEKT 60
    |||||
Dd 1 MSPARLPRHLFCVLVLLLVPAAMCGPGRVVGSRRRPRKLVPLAYKQFSPNVPEKT 60
    |||||
QY 61 LGASGRREGKTAARSERFKELTPTNPNPDIIFKDEENTGADRLMTQRCDRNLNSLAISVMN 120
    |||||
Dd 61 LGASGRREGKTAARSERFKELTPTNPNPDIIFKDEENTGADRLMTQRCDRNLNSLAISVMN 120
    |||||
QY 121 QMPGVKRLVTEGMDGDHSEESLHYEGRAVDITTSDBDRKKYGLRLARLAEAGFDWYY 180
    |||||
Dd 121 QMPGVKRLVTEGMDGDHSEESLHYEGRAVDITTSDBDRKKYGLRLARLAEAGFDWYY 180
    |||||
QY 181 ESKAHVHCYSKSEHSAAKTGCGCPAGAOVRLSEGARVALSAVRPGDVLAMGEDGSPTE 240
    |||||
Dd 181 ESKAHVHCYSKSEHSAAKTGCGCPAGAOVRLSEGARVALSAVRPGDVLAMGEDGSPTE 240
    |||||
QY 241 SDVLIIFLDREPHRLRAFOVLETODPPRRLLTPAHLLEFTADNHTEPARFRATASHVQ 300
    |||||
Dd 241 SDVLIIFLDREPHRLRAFOVLETODPPRRLLTPAHLLEFTADNHTEPARFRATASHVQ 300
    |||||
QY 301 GOYLVAGVPGLOPARAAVSTHVALGAYAPLTKHGLTVEDVYASCFPAADHHLAQLA 360
    |||||
Dd 301 GOYLVAGVPGLOPARAAVSTHVALGAYAPLTKHGLTVEDVYASCFPAADHHLAQLA 360
    |||||
QY 361 FWPRLRFLHSLAMGSWTPGEGVHWYPOLLYRLGRLLLEGGSFHPLMGSGAGS 411
    |||||
Dd 361 FWPRLRFLHSLAMGSWTPGEGVHWYPOLLYRLGRLLLEGGSFHPLMGSGAGS 411
    |||||
```

RESULT 5
W97767 standard; Protein; 411 AA.
AC W97767;
DE 21-MAY-1999 (first entry)
DE Mouse Indian hedgehog (Ihh) protein.
KW Indian hedgehog; Ihn protein; mouse; dopaminergic; GABA-ergic;
KW ptc therapeutic; patched; signal transduction; Parkinson's disease;
KW Huntington's disease; amyotrophic lateral sclerosis;
KW cerebral ischaemia; hypoxia; neuroprotective; therapy.
OS Mus sp.
PN WO9904775-A2.
PD 04-FEB-1999.
PE 24-JUL-1998; U15419.
PR 24-JUL-1997; US-900220.
PA (ONTO-) ONTOGENY INC.
PI Mahanthappa NK, Miao N, Pang K, Wang M;
DR N-PSDB; X07273.
PT Increasing the survival of neuronal, dopaminergic and GABA-ergic
PT cells - by using a ptc therapeutic such as a protein kinase
PT inhibitor, or an agent derived from hedgehog polypeptides, useful in
PT the treatment of Parkinson's disease
PS Disclosure; Page 87-88; 138pp; English.
CC This polypeptide is mouse Ihn Indian hedgehog protein. The
CC invention is based on the finding that hedgehog proteins are useful
CC as protective agents in the treatment and prophylaxis of
CC neurodegenerative disorders resulting from the loss of dopaminergic
CC and/or GABA-ergic neurons, or the general loss of tissue from the
CC substantia nigra. Exemplary disorders include Parkinson's disease,
CC Huntington's disease (both claimed), amyotrophic lateral sclerosis
CC and cerebral ischaemia. The invention relates to hedgehog
CC therapeutics (i.e. hedgehog polypeptides and gene therapy constructs
CC e.g. constructs encoding recombinant hedgehog polypeptides and
CC trans-activation constructs for altering hedgehog gene regulatory
CC sequences) and ptc therapeutics (i.e. agents which mimic the effect
CC of naturally occurring hedgehog proteins on patched signalling)

CC that are effective in both human and animal subjects. Human Ihh
 CC and Dhh polypeptides (see W97/63-64) are preferred. The products
 CC can also be used for the maintenance of differentiated neurons in
 CC cultures, and to enhance the implantation of such neuronal cells in
 CC an animal. They can be used to prevent or treat neurodegenerative
 CC conditions arising from the use of certain drugs, and in the
 CC prevention and/or treatment of hypoxia, e.g. as a neuroprotective
 CC agent.
 SO Sequence 411 AA;

Query Match 93.8%; Score 2041; DB 1; Length 411;
 Best Local Similarity 94.6%; Pred. No. 35-200;
 Matches 389; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 1 MSPARLRRLHFCVLVLLLVPAAMGCGPGRVGSRRPRPKLYPLAYKQFSPNVEKT 60
 DB 1 MSPARLRRLHFCVLVLLLVPAAMGCGPGRVGSRRPRPKLYPLAYKQFSPNVEKT 60
 QY 61 LGASGRYEGKTAASSEKELTPNYNDIIFKDEENTGADRLMTORCKDRLSLAISVMN 120
 DB 61 LGASGRYEGKTAASSEKELTPNYNDIIFKDEENTGADRLMTORCKDRLSLAISVMN 120
 QY 121 QMPVKRLVTEGMDGDHSESLHYEGRAVDITSDNRNRYGLARLAVAGFDWY 180
 DB 121 QMPVKRLVTEGMDGDHSESLHYEGRAVDITSDNRNRYGLARLAVAGFDWY 180
 QY 181 ESKAHVCSVSESAAKTGCGCPAGAOVRLSEGAVALSAVRPGDRILAMGEGSPTE 240
 DB 181 ESKAHVCSVSESAAKTGCGCPAGAOVRLSEGAVALSAVRPGDRILAMGEGSPTE 240
 QY 241 SDVLIIFLDREPRRLRAFOVETODPPRLALTPAHLFTADNHTPEAPRAFASHVQ 300
 DB 241 SDVLIIFLDREPRRLRAFOVETODPPRLALTPAHLFTADNHTPEAPRAFASHVQ 300
 QY 301 GOYLVAGVGPGLQPARAAVSTHVALGSAVPLTRHGTLVEDVVASCPAAVADHHLAQLA 360
 DB 301 GOYLVAGVGPGLQPARAAVSTHVALGSAVPLTRHGTLVEDVVASCPAAVADHHLAQLA 360
 QY 361 FWPLRLFSLAMGSWTPGEGVHWYPOLLYRLGRLLLEGSFHPGLMSGAGS 411
 DB 361 FWPLRLFSLAMGSWTPGEGVHWYPOLLYRLGRLLLEGSFHPGLMSGAGS 411

RESULT 6
 Y05512 ID Y05512 standard; Protein; 411 AA.
 AC Y05512;
 DT 05-JUL-1999 (first entry)
 DE Mouse Indian hedgehog protein Ihh.
 KW Indian hedgehog; Ihh protein; mouse; hedgehog therapeutic;
 KM ptc therapeutic; patched; signal transduction; muscle atrophy;
 OS cecelia; muscular myopathy; myoblastic sarcoma; therapy.
 MS Mus sp.
 PN MO9910004-A2.
 PD 04-MAR-1999.
 PF 28-AUG-1998; U17922.
 PR 29-AUG-1997; US-057394.
 PA (ONTO-) ONTOGENY INC.
 PI Bladen CS, Currie PD, Hughes SM, Ingham PW;
 DR WPI: 99-243557/20.
 DR N-PSDB: X25100.
 PT A new method to regulate muscle growth
 PS Disclosure; Page 114-115; 130pp; English.
 CC The present sequence is mouse Indian hedgehog protein Ihh. The
 CC invention relates to a method for modulating the formation and/or
 CC maintenance of muscle tissue by ectopically contacting muscle
 CC cells, especially muscle stem/progenitor cells, in vitro or in
 CC vivo, with a hedgehog therapeutic (i.e. hedgehog polypeptides and
 CC gene therapy constructs) or ptc therapeutic (i.e. a small organic
 CC molecule that mimics the effect of hedgehog proteins on patched
 CC signaling, or activates or potentiates patched signaling) in an
 CC amount effective to alter the growth state of the treated cells.

CC Also claimed is a method for treatment or prevention of disorders
 CC of, or surgical or cosmetic repair of, such muscle tissues by
 CC administering a hedgehog polypeptide or ptc therapeutic. The
 CC disorder may be muscle atrophy, in particular skeletal muscle
 CC atrophy or cardiac muscle atrophy, cachexia, or muscular myopathy
 CC (all claimed). The hedgehog polypeptide or ptc therapeutic can
 CC inhibit growth of myoblastic-derived tissue to provide treatment of
 CC myoblastic or neoplastic growth of muscle tissue such as in
 CC myoblastic sarcoma (also claimed). The hedgehog therapeutic
 CC preferably comprises at least a bioactive extracellular portion of
 CC a hedgehog protein (see Y05510-19) encoded by a vertebrate hedgehog
 CC gene (see X25098-107), especially a human hedgehog gene.
 SO Sequence 411 AA;

Query Match 93.8%; Score 2041; DB 1; Length 411;
 Best Local Similarity 94.6%; Pred. No. 35-200;
 Matches 389; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 1 MSPARLRRLHFCVLVLLLVPAAMGCGPGRVGSRRPRPKLYPLAYKQFSPNVEKT 60
 DB 1 MSPARLRRLHFCVLVLLLVPAAMGCGPGRVGSRRPRPKLYPLAYKQFSPNVEKT 60
 QY 61 LGASGRYEGKTAASSEKELTPNYNDIIFKDEENTGADRLMTORCKDRLSLAISVMN 120
 DB 61 LGASGRYEGKTAASSEKELTPNYNDIIFKDEENTGADRLMTORCKDRLSLAISVMN 120
 QY 121 QMPVKRLVTEGMDGDHSESLHYEGRAVDITSDNRNRYGLARLAVAGFDWY 180
 DB 121 QMPVKRLVTEGMDGDHSESLHYEGRAVDITSDNRNRYGLARLAVAGFDWY 180
 QY 181 ESKAHVCSVSESAAKTGCGCPAGAOVRLSEGAVALSAVRPGDRILAMGEGSPTE 240
 DB 181 ESKAHVCSVSESAAKTGCGCPAGAOVRLSEGAVALSAVRPGDRILAMGEGSPTE 240
 QY 241 SDVLIIFLDREPRRLRAFOVETODPPRLALTPAHLFTADNHTPEAPRAFASHVQ 300
 DB 241 SDVLIIFLDREPRRLRAFOVETODPPRLALTPAHLFTADNHTPEAPRAFASHVQ 300
 QY 301 GOYLVAGVGPGLQPARAAVSTHVALGSAVPLTRHGTLVEDVVASCPAAVADHHLAQLA 360
 DB 301 GOYLVAGVGPGLQPARAAVSTHVALGSAVPLTRHGTLVEDVVASCPAAVADHHLAQLA 360
 QY 361 FWPLRLFSLAMGSWTPGEGVHWYPOLLYRLGRLLLEGSFHPGLMSGAGS 411
 DB 361 FWPLRLFSLAMGSWTPGEGVHWYPOLLYRLGRLLLEGSFHPGLMSGAGS 411

RESULT 7
 R77343 ID R77343 standard; Protein; 336 AA.
 AC R77343;
 DT 18-MAR-1996 (first entry)
 DE Mouse Indian hedgehog protein.
 KW Mouse; Indian hedgehog protein; probe; primer; diagnostic;
 KM nervous system disorder; gene therapy; antibody.
 OS Mus musculus.
 PN MO9518856-A1.
 PD 13-JUL-1995.
 PF 30-DEC-1994; U14992.
 PR 30-DEC-1993; US-176427.
 PR 14-DEC-1994; US-356060.
 PA (HARD) HARVARD COLLEGE.
 PI (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
 PI Ingham PW, McMahon AB, Tabin CJ;
 DR WPI: 95-255060/33.
 DR N-PSDB: Q91640.
 PT Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful
 PT to treat degenerative nervous system disorder(s) and in gene
 PT therapy.
 PS Claim 17; Page 137-38; 210pp; English.
 CC The sequence represents a mouse Indian hedgehog protein, homologous
 CC to a Drosophila hedgehog protein (R77337), and is encoded by a cDNA

CC isolated by low stringency screening of a mouse genome DNA library
CC and screening of an 8.5 day post coitum cDNA library. Probes and
CC primers derived from hedgehog sequences may be used as diagnostic
CC agents for neuromuscular, autonomic or central nervous system
CC disorders, and the gene may also be used in gene therapy.
CC Antibodies generated from the protein may be used as therapeutic or
CC research reagents.
SQ Sequence 336 AA;

Query Match 77.5%; Score 1686; DB 1; Length 336;
Best Local Similarity 94.9%; Pred. No. 6.6e-164;
Matches 319; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 76 ERFKELPNYNDIIFKDEENTGADRLMTORCKRLNSLAISVMQMPGVKLTVEGMD 135
DB 1 ERFKELPNYNDIIFKDEENTGADRLMTORCKRLNSLAISVMQMPGVKLTVEGMD 60
QY 136 DGHSESLHTEGRAVDITTSDBRNKYGGLARLAVENGFPWYVESKAHVCSKSEHS 195
DB 61 DGHSESLHTEGRAVDITTSDBRNKYGGLARLAVENGFPWYVESKAHVCSKSEHS 120
QY 196 AAKTGGCFPAQVRLSESGARVALSAVRPGDVLAMGEDSGPTSDVLIIFDRPHRLR 255
DB 121 AAKTGGCFPAQVRLSESGARVALSAVRPGDVLAMGEDSGPTSDVLIIFDRPHRLR 180
QY 256 AFQVETODPPRRLLTPAHLIFTDNHTEPARFRATFASHVQPGQYVLAVGPGLOPA 315
DB 181 AFQVETODPPRRLLTPAHLIFTDNHTEPARFRATFASHVQPGQYVLAVGPGLOPA 240
QY 316 RVAANSTHVALGAVAPLRFKHTLVVEDYVASCFAAVADHHLAQLAFWPLRFLHSLAMGSM 375
DB 241 RVAANSTHVALGAVAPLRFKHTLVVEDYVASCFAAVADHHLAQLAFWPLRFLHSLAMGSM 300
QY 376 TPGECHMYPOLLYRLGRLLLESGSFHPLGMSGAGS 411
DB 301 TPGECHMYPOLLYRLGRLLLESGSFHPLGMSGAGS 336

RESULT 8
R77344
ID R77344 standard; Protein; 312 AA.
AC R77344:
DE 18-MAR-1996 (first entry)
DT Human Indian hedgehog protein.
KW Human; Indian hedgehog protein; probe; primer; diagnostic;
KW nervous system disorder; gene therapy; antibody.
OS Homo sapiens.
PN M09518856-A1.
PD 13-JUL-1995.
PF 30-DEC-1994; U14992.
PR 30-DEC-1993; US-176427.
PR 14-DEC-1994; US-356060.
PA (HARD) HARVARD COLLEGE.
PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
PI Ingham PW, Mcmanan AP, Tabin CJ;
DR WPI: 95-255060/33.
DR N-PSDB: 091641.
PT Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful
PT to treat degenerative nervous system disorder(s) and in gene
PT therapy.
PS Claim 17; Page 146-47; 210pp; English.
CC The sequence represents a mouse Indian hedgehog protein, homologous
CC to a Drosophila hedgehog protein (R77337), and is encoded by a cDNA
CC isolated by screening of a human fetal lung 5'-stretch plus cDNA
CC library. Probes and primers derived from hedgehog sequences may be
CC used as diagnostic agents for neuromuscular, autonomic or central
CC nervous system disorders, and the gene may also be used in gene
CC therapy. Antibodies generated from the protein may be used as
CC therapeutic or research reagents.
SQ Sequence 312 AA;

Query Match 75.7%; Score 1646; DB 1; Length 312;
Best Local Similarity 100.0%; Pred. No. 7.1e-160;
Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 RLMTORCKDRNLNSLAISVMQMPGVKLTVEGMDGHSESLHTEGRAVDITTSDBR 160
DB 101 RLMTORCKDRNLNSLAISVMQMPGVKLTVEGMDGHSESLHTEGRAVDITTSDBR 61
QY 161 NKYGGLARLAVENGFPWYVESKAHVCSKSESAAKTGGCFPAQVRLSESGARVAL 220
DB 62 NKYGGLARLAVENGFPWYVESKAHVCSKSESAAKTGGCFPAQVRLSESGARVAL 121
QY 221 SAVRPGDVLAMGEDSGPTSDVLIIFDRPHRLRAVITITODPPRRLLTPAHLFTA 280
DB 122 SAVRPGDVLAMGEDSGPTSDVLIIFDRPHRLRAVITITODPPRRLLTPAHLFTA 181
QY 281 DNHTEPARFRATFASHVQPGQYVLAVGPGLOPARVAASTHVALGAYADLTGHTLVV 340
DB 182 DNHTEPARFRATFASHVQPGQYVLAVGPGLOPARVAASTHVALGAYADLTGHTLVV 241
QY 341 EDVYASCPAAVADHHLAQLAFWPLRFLHSLAMGSMTPGECHMYPOLLYRLGRLLLEGS 400
DB 242 EDVYASCPAAVADHHLAQLAFWPLRFLHSLAMGSMTPGECHMYPOLLYRLGRLLLEGS 301
QY 401 FHPGLMSGAGS 411
DB 302 FHPGLMSGAGS 312

RESULT 9
W61488
ID W61488 standard; Protein; 437 AA.
AC W61488:
DE 20-OCT-1998 (first entry)
DT Mouse sonic hedgehog (shh) protein.
KW Hedgehog polypeptide; neuronal cell proliferation; zebrafish; shh;
KW cholesterol biosynthesis; pituitary gland gene expression; mouse;
KW liggy-winkle hedgehog protein; twih; sonic hedgehog; major limb trauma;
KW neuronal degeneration; nerve-sparing agent; Smith-Lemli-Opitz syndrome.
OS Mus sp.
PN M09830576-A1.
PD 16-JUL-1998.
PF 07-OCT-1997; U15753.
PR 02-OCT-1997; US-061323.
PR 07-OCT-1996; US-729743.
PA (UYUO) UNIV JOHNS HOPKINS SCHOOL OF MEDICINE.
PI Beachy PA, Porter JA;
DR WPI: 98-399053/34.
PT New hedgehog-derived polypeptide(s) - used to develop products for
PT modulating proliferation or differentiation of neuronal cells,
PT cholesterol biosynthesis or transport or expression of pituitary
PT gland gene(s)
PS Claim 13; Pages 153-154; 210pp; English.
CC This represents a mouse sonic hedgehog (shh) protein sequence. The
CC invention provides methods and compounds for modulating proliferation or
CC differentiation of neuronal cells, cholesterol biosynthesis or transport
CC or expression of pituitary gland genes. The method for affecting
CC cholesterol biosynthesis or transport in a cell comprises contacting a
CC cell with a compound that affects hedgehog, thereby affecting cholesterol
CC biosynthesis or transport. The methods for inhibiting the neural inducing
CC activity of a hedgehog polypeptide in cells, and for inducing pituitary
CC gland gene expression utilizes sequences selected from a zebrafish
CC liggy-winkle hedgehog (twih) protein (W61485), a zebrafish sonic hedgehog
CC (shh) protein (W61486), a chicken shh protein (W61487) or a mouse shh
CC protein (W61488). The products and methods provide for compounds which
CC can affect hedgehog activity. They can be used for treating disorders
CC which arise from neuronal degeneration or abnormal function. They can
CC also be used as nerve-sparing agents or in restoring or promoting
CC appropriate patterning during the healing of major limb trauma. They can
CC also be used for treating Smith-Lemli-Opitz syndrome. The products can
CC also be used for detection and diagnosis.
SQ Sequence 437 AA;

Query Match	58.8%	Score 1279.5;	DB 1;	Length 437;
Best Local Similarity	61.1%;	Pred. No. 3e-122;		
Matches 262;	Conservative 39;	Mismatches 91;	Indels 37;	Gaps 8

Qy	13	CLVLL-----LLVPAAMCGGGRVYSGRRPRKLYPLATQKQFSPNVEKTLASGRYEG	69
Db	8	CFVLVASSLLVCPQL-ACGFGRGG-KRRHKPLTPLYAQFIPNNAEKTLASGRYEG	65
Qy	70	KIASSEFFKELTPYNDIIFKDEENTGADRLMTORCKDRLNSLAISVMNQMGVKRLV	129
Db	66	KITNSEFFKELTPYNDIIFKDEENTGADRLMTORCKDRLNSLAISVMNQMGVKRLV	125
Qy	130	TEGMEDEHHSEESIHYEGRAVDITTSRDNRKTKGLLARLVEAGFPMVYSESAHYHCS	189
Db	126	TEGMEDEHHSEESIHYEGRAVDITTSRDNRKTKGLLARLVEAGFPMVYSESAHYHCS	185
Qy	190	VKSEHSAAKATGGCGPAGAQVRLSEGAVALSAVRPGDRLVAMEGDSPTFSDVLIPLDR	249
Db	186	VKAENSAVAKSGGCP- Q SATVHLEGGCKLVKDLRPGDRLVLAADQGRLLYSDPLFLDR	245
Qy	250	EPHRLRAQVLETQDPPRLALTPLHLLFTL-DNHTEPARFRTAASHVOPGOYVLVAG	308
Db	246	DEGAKRYEYVLELEPERRLLTAAHLFTVAPHNDSGETPGSALFASRPVRCGRVYVA	305
Qy	309	YPG-----LQPAEVAVS-THVALGAVAPLTGHTLVYDVVASQFAVAADHNLQALEMP	363
Db	306	ERGGCRLLLPAAVHSVTLREEDAGATAPLTHAGTILLIRVLASQYANVIEESVAMHRAFAP	365
Qy	364	LRLFHSI-----ANGSWTPGEGVMYDQLLYRLGLLLEE	398
Db	366	FRLLAALLAALAPARTDGGGGSIPAAQSATEARGA-EPTAGIMYQDLLYHISTWLDS	424
Qy	399	GSFHLGMS 407	
Db	425	ETMHPLGMA 433	

RESULT	10
R77339	
ID	R77339 standard; Protein; 437 AA.
AC	R77339;
DT	08-MAR-1996 (first entry)
DE	Mouse sonic hedgehog protein.
KW	Mouse; sonic hedgehog protein; probe; primer; diagnostic;
KW	nervous system disorder; gene therapy; antibody.
OS	Mus musculus.
FT	Key
FT	Location/Qualifiers
FT	peptide
FT	1..24
FT	/note="signal peptide"
FT	25..30
FT	/note="conserved sequence (R77349)"
FT	279
FT	/note="N-linked glycosylation site"
FT	
PN	W09518856-A1.
PD	13-JUL-1995.
PF	30-DEC-1994; U14992.
PR	30-DEC-1993; U5-176427.
PR	14-DEC-1994; U5-356060.
PA	(HARD) HARVARD COLLEGE.
PA	(IMCR) IMPERIAL CANCER RES TECHNOLOGY.
PI	Ingham PM, McMahon AP, Tabin CJ;
DR	WPI; 95-255060/33.
DR	N-PSDB; 091637.
PT	Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful to treat degenerative nervous system disorder(s) and in gene therapy.
PS	Claim 17; Page 139-41; 210pp; English.
CC	The sequence represents a mouse sonic hedgehog protein, homologous to a Drosophila hedgehog protein (R77337), and is encoded by a cDNA isolated from an 8.5-day post coitum mouse CC library. Probes and primers derived from the sonic hedgehog gene may be used as diagnostic agents for neuromuscular, autonomic or

CC central nervous system disorders, and the gene may also be used in
CC therapy. Antibodies generated from the protein may be used
CC as therapeutic or research reagents.
SQ Sequence 437 AA;

Query Match Similarity	58.7%	Score 1276.5	DB 1	Length 437
Best Local Similarity	60.8%	Pred. No. 6e-122		
Matches	261	Conservative	40	Mismatches 91; Indels 37; Gaps
QY	13	CLVLL---LTVYPAMCGGPERVVGSRRRPRKLVPLAYKQSPVNPVKTLGASGRYEG	69	
Db	8	CFVLVLASLLCPGL-ACGPERGFG-KRRHPRKLLPLAYKQFIPVNAKTLTGASGRYEG	65	
QY	70	KIARSSERPEKLTPLYNPNDIIFKDEENTGADLMTORCDRLNSLAISSYMNOPVYKLEY	129	
Db	66	KTRNSERPEKLTPLYNPNDIIFKDEENTGADLMTORCDKLNALAISSYMNOPVYKLEY	125	
QY	130	TEGWDEDHGHSSESLHYEGRAVDITTSDBRNRKYLGLARLAVEAGFDWVYYESKAHVHCS	189	
Db	126	TEGWEDDHGHSSESLHYEGRAVDITTSDBRNRKYLGLARLAVEAGFDWVYYESKAHVHCS	185	
QY	130	VKSEHSAAKTGCGCPAGQVRLSEGARALSAVPRGDVYLA MGDEGSPTEFDYLFLDR	249	
Db	136	VKAENSVAAKSGGCGPSATVYHLDEGGTKLYDLTRGDDVYLAADOGRLLYSDFTFLDR	245	
QY	250	EPHRLRAFOVLETODPRRLALTPRHLLFTA-DNTEPRAFRATFASHVQGOYVLVAG	308	
Db	246	DEGAKKVFYETLEPRERLLTALHLLFVAHNDSGPTPGPSALFASRVRRGQRYUYVA	305	
QY	309	VPG-----IQPARVAVS-THVALGAYARPLTKGTLVEDVYVASCFAAVADHHLAQALFAP	363	
Db	306	ERGGDRLLPAVAHVSVTLREEPAGAYAPLTAGTLLINVLNASCVAIEESHMAHAFAP	365	
QY	364	LRLFHSL-----AMGSWTPEGVHWYPOLLYRIGRLLE	398	
Db	366	FLRLAHLAALAPARTDDGGGGSIPAQASATARGA-EPTAGIHMYSQLYHIGTGLDS	424	
QY	399	GSFHLIGMS 407		
Db	425	ETNHPLGMA 433		

RESULT 11
ID W94471 standard: Protein; 437 AA.
AC W94471;
DT 29-APR-1999 (first entry)
DE Mouse Shh hedgehog protein sequence.
KW Patched; hedgehog; ptc therapeutic; neuroprotective; neuronal cell;
KW brain infarction; cerebral infarction; transient ischaemic attack;
KW stroke; cerebral infarct volume; spinal cord; oedema; trauma;
KW haemorrhage; encephalomyelitis; coronary bypass; cerebral hypoxia.
OS Mus sp.
PN WC0900117-A2.
PD 07-JAN-1999.
PF 26-JUN-1998; U13387.
PR 27-JUN-1997; US-883656.
PA (ONTO-) ONTOGENY INC.
PI Mahantappa, NK;
DR WPI: 99-095458/08.
DR N-PSDS; X16185.
PT Method for limiting damage to neurons caused by ischaemic or epoxic
PT conditions - is used for the treatment and prevention of e.g.
PT cerebral infarction, stroke and transient ischaemic attacks
CC Disclosure; page 68-70; 104pp; English.
CC A method has been developed for limiting the damage to neuronal cells by
CC ischaemic or epoxic conditions by administering a ptc (patched)
CC therapeutic agent to reduce cerebral infarct volume (CIV). Damage to
CC neuronal cells can also be limited by administering a gene activation
CC construct which recombines with the genomic hedgehog gene to provide a
CC heterologous transcription regulator linked to the coding region of this
CC gene. Administration of the ptc therapeutic agent is used to protect

PN WO910004-A2.
PD 04-MAR-1999.
PR 28-AUG-1998; U17922.
PR 29-AUG-1997; US-057394.
PI (ONTO-) ONTOGENY INC.
PI Bladgen CS, Currie PD, Hughes SM, Ingham PW;
DR WPI: 99-24357/20.
DR N-PSDB: X25101.
PT A new method to regulate muscle growth.
PS Disclosure: Page 115-116; 130pp; English.
CC The present sequence is mouse Sonic hedgehog protein Shh. The
CC invention relates to a method for modulating the formation and/or
CC maintenance of muscle tissue by ectopically contracting muscle
CC cells, especially muscle stem/progenitor cells, in vitro or in
CC vivo, with a hedgehog therapeutic (i.e. hedgehog polypeptides and
CC gene therapy constructs) or ptc therapeutic (i.e. a small organic
CC molecule that mimics the effect of hedgehog proteins on patched
CC signaling, or activates or potentiates patched signaling) in an
CC amount effective to alter the growth state of the treated cells.
CC Also claimed is a method for treatment or prevention of disorders
CC of, or surgical or cosmetic repair of, such muscle tissues, by
CC administering a hedgehog polypeptide or ptc therapeutic. The
CC disorder may be muscle atrophy, in particular skeletal muscle
CC atrophy or cardiac muscle atrophy, cachexia, or muscular myopathy
CC (all claimed). The hedgehog polypeptide or ptc therapeutic can
CC inhibit growth of myoblastic-derived tissue to provide treatment of
CC hyperplastic or neoplastic growth of muscle tissue such as in
CC myoblastic sarcoma (also claimed). The hedgehog therapeutic
CC preferably comprises at least a bioactive extracellular portion of
CC a hedgehog protein (see Y0510-19) encoded by a vertebrate hedgehog
CC gene (see X25098-107), especially a human hedgehog gene.
SO Sequence 437 AA:

Query Match 58.7%; Score 1276.5; DB 1; Length 437;
Best Local Similarity 60.8%; Pred. No. 6e-122;
Matches 261; Conservative 40; Mismatches 91; Indels 37; Gaps 8;

13 CLVLL-LLVPAAMGCGPRGVGSRPRKLVPLAYKQFSPVNPVKTKGASGRYG 69
8 CGLVLIASSLLVCPGL-ACGPGRGFC-KRRHPKKLPLAYKQFIPVNAKTKGASGRYEG 65
70 KARSSEKELTPVNPDIIFKDEENTGADRLMTORCKDRNLNSLAISVMNOMPGLRV 129
66 KTRNSERKELTPVNPDIIFKDEENTGADRLMTORCKDRNLNSLAISVMNOMPGLRV 125
130 TEGWDEGHSESLHYEGRAVDITTSDDRNKYGGLARLAVEAGFDWYYSKAHVCS 189
126 TEGWDEGHSESLHYEGRAVDITTSDDRNKYGGLARLAVEAGFDWYYSKAHVCS 185
190 VSESSAAKTCGCPAGAOVRLSGARVALSAVRGDLVLMGEGSTFSDVLTFLDR 249
186 VAEKSVAAKSGCGPAGVTLHLEGGCTLVLDLRGDDVLAADOGRLYDFTFLDR 245
250 EPHRAFOVITOPPRRLATLPAHLFTA-DNHTEPARFRATFASHVOGQYVAVG 308
246 DSGAKVFTVITLTPRELLTLAHLFVAPRNDGSPFGSALFASVVRGQRYVYA 305
309 VEG---LQPARVAANS-THVALGAVAPLTKHGTLVEDVVASCPAAVADHHLAQLAFWP 363
306 ERGGDRRLPAVAVSHVTLKEEAGAVAPLTAHGTILINVLASCAVIEHSMARAFAP 365
364 LRLFLSL-----AAGSWTPGEGVHYTPOLLVYLGRLLEE 398
366 FLTAHLAALAPARTDGGGSGIPAASQATEARGH-EPTAGIHVYSOLLHYIGTWLDS 424
399 GSEHPLGMS 407
425 ETMHPGLMA 433

RESULT 14
W61486

ID W61486 standard; Protein; 418 AA.
AC W61486;
DE 20-OCT-1998 (first entry)
DT Zebrafish sonic hedgehog (shh) protein.
KW Hedgehog polypeptide; neuronal cell proliferation; zebrafish;
KW Cholesterol biosynthesis; pituitary gland gene expression; shh;
KW tiggly-winkle hedgehog protein; twhh; sonic hedgehog; major limb trauma;
KW neuronal degeneration; nerve-sparing agent; Smith-Lemli-Opitz syndrome.
OS Brachydanio rerio.
PN WO9830576-A1.
PD 16-JUL-1998.
PR 07-OCT-1997; U15753.
PR 02-OCT-1997; US-061323.
PR 07-OCT-1996; US-729743.
PI (UNIV) UNIV JOHNS HOPKINS SCHOOL OF MEDICINE.
PI Beachy PA, Porter JA;
DR WPI: 98-399053/34.
PT New hedgehog-derived polypeptide(s) - used to develop products for
PT modulating proliferation or differentiation of neuronal cells,
PT cholesterol biosynthesis or transport or expression of pituitary
PT gland gene(s)
PS Claim 13: Pages 150-151; 210pp; English.
CC This represents a zebrafish sonic hedgehog (shh) protein sequence. The
CC invention provides methods and compounds for modulating proliferation or
CC differentiation of neuronal cells, cholesterol biosynthesis or transport
CC or expression of pituitary gland genes. The method for affecting
CC cholesterol biosynthesis or transport in a cell comprises contacting a
CC cell with a compound that affects hedgehog, thereby affecting cholesterol
CC biosynthesis or transport. The methods for inhibiting the neural inducing
CC activity of a hedgehog polypeptide in cells, and for inducing pituitary
CC gland gene expression utilizes sequences selected from a zebrafish
CC tiggly-winkle hedgehog (twhh) protein (W61485), a zebrafish sonic hedgehog
CC (shh) protein (W61486), a chicken shh protein (W61487) or a mouse shh
CC protein (W61488). The products and methods provide for compounds which
CC can affect hedgehog activity. They can be used for treating disorders
CC which arise from neuronal degeneration or abnormal function. They can
CC also be used as nerve-sparing agents or in restoring or promoting
CC appropriate patterning during the healing of major limb trauma. They can
CC also be used for treating Smith-Lemli-Opitz syndrome. The products can
CC be used for detection and diagnosis.
SO Sequence 418 AA:

Query Match 58.6%; Score 1274; DB 1; Length 418;
Best Local Similarity 61.2%; Pred. No. 1e-121;
Matches 254; Conservative 46; Mismatches 101; Indels 14; Gaps 4;

5 RLRLPHCLVLLLLVPAAMGCGPRGVGSRPRKLVPLAYKQFSPVNPVKTKGAS 64
2 RLRLVLLVSLTSLVY-SGLACGPRGFG-KRRHPKKLPLAYKQFIPVNAKTKGAS 59
65 GRYECKIARSSERKELTPVNPDIIFKDEENTGADRLMTORCKDRNLNSLAISVMNOMP 124
60 GRYECKIARSSERKELTPVNPDIIFKDEENTGADRLMTORCKDRNLNSLAISVMNOMP 119
125 VKLRTBEGWDEGHSESLHYEGRAVDITTSDDRNKYGGLARLAVEAGFDWYYSKA 184
120 VKLRTBEGWDEGHSESLHYEGRAVDITTSDDRNKYGGLARLAVEAGFDWYYSKA 179
185 HVHGSVSESSAAKTCGCPAGAOVRLSGARVALSAVRGDLVLMGEGSTFSDVLT 244
180 HIRCSVKAENSVAAKSGCGPAGVTLHLEGGCTLVLDLRGDDVLAADOGRLYDFTFLDR 239
245 IFDREPHRLAFOVITOPPRRLATLPAHLFTA-DNHTEPARFRATFASHVOGQYV 304
240 METDRDSTTRRVFTVIEQPEVKITLPAHLFVLDNSTEDLTMAAAYSSVBAQKV 299
305 LVAGVPG-LQPARVAANS-THVALGAVAPLTKHGTLVEDVVASCPAAVADHHLAQLAFWP 363
300 MVVDDSGOLKSVIVORTTEORGSFAPVTAHGTVYVDRILASCAVIEHSMARAFAP 359
364 LRLFL-----SLAGSWTPGEGVHYTPOLLVYLGRLLEEBSFHLGMS 407

Db 360 ARLYTVSSFLPONSRRSNATLQOEGVHWYSRLLYQMGTWLDSNMHLPLGMS 414

RESULT 15

ID R77341 standard; Protein; 475 AA.

AC R77341.

DT 14-MAR-1996 (first entry)

DE Human sonic hedgehog protein.

KW Human; sonic hedgehog protein; probe; primer; diagnostic;

OS nervous system disorder; gene therapy; antibody.

FT Key Location/Qualifiers

FT peptide 24..29

FT misc_difference 463 /note= "conserved sequence (R77349)"

FT /note= "unspecified amino acid"

PN W09518856-A1.

PD 13-JUL-1995.

PR 30-DEC-1994; U14992.

PR 30-DEC-1993; US-176427.

PR 14-DEC-1994; US-356060.

PA (HARD) HARVARD COLLEGE.

PI (IMCR) IMPERIAL CANCER RES TECHNOLOGY.

PI Ingham PW, McMahon AP, Tabin CJ;

DR N-PSDB; Q91639.

DR Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful

PT to treat degenerative nervous system disorder(s) and in gene

PT therapy.

PS Claim 17; Page 143-45; 210pp; English.

CC The sequence represents a human sonic hedgehog protein, homologous

CC to Drosophila hedgehog protein (R77337), and is encoded by a cDNA

CC isolated from a human fetal lung cDNA library. Probes and primers

CC derived from the sonic hedgehog gene may be used as diagnostic

CC agents for neuromuscular, autonomic or central nervous system

CC disorders, and the gene may also be used in gene therapy.

CC Antibodies generated from the protein may be used as therapeutic or

CC research reagents.

SC Sequence 475 AA:

Query Match 58.4%; Score 1271; DB 1; Length 475;

Best Local Similarity 57.7%; Pred. No. 2.5e-121;

Matches 267; Conservative 45; Mismatches 85; Indels 66; Gaps 10;

QY 13 CTVLLLT--LVYPAMGCCPGRVVGSRRRPRKLVPLAYKQSPNYPEKTIGASGRYEKG 70

DB 7 CLLLVVSSLSLVCSGLACGPRGFG-KRRHPKRLPLAYKQFIPNVAEKTIGASGRYEKG 65

QY 71 IARSSERFRELTPNYPNDIIFKDEENTGADRLMTORCKDRLNSLAISVMNOMPVGLRYT 130

DB 66 ISRNSERFRELTPNYPNDIIFKDEENTGADRLMTORCKDRLNSLAISVMNOMPVGLRYT 125

QY 131 EGMDEDGHHSESLAYEGRAVDITTSDRDRNKYGLLARLAVEAGFDWYVESKAHVCSV 190

DB 126 EGMDEDGHHSESLAYEGRAVDITTSDRDRNKYGLLARLAVEAGFDWYVESKAHVCSV 185

QY 191 KSEHSAATACTGCGFPAGAOVRLSEGARVALSAVRPDRLVAMGEDSPTFSDVLIFLDRE 250

DB 186 KAENSVAAKSGCGFPAGAOVRLSEGARVALSAVRPDRLVAMGEDSPTFSDVLIFLDRE 245

QY 251 PHRLFAFYIETODPPRLALTPAHLFTA--DNHT-EPARF-----RATF 294

DB 246 DGAKKFTYIETREPPRLILTAHLLETPAHPNDSATGEPEASSGSGPBGALGPRLALF 305

QY 295 ASHVDPQGVYLVV--GVPGLOPARVAVS-THVALGAYAPLTKHGLTVEDVVAISCF 349

DB 306 ASRVRRGQGVYVVAERDGRRLPAVHVSYTLSEEAAGAYAPLTAQGTILINRVLASCYA 365

QY 350 AVADHLLAQLAFPLRLFHSLSL-----AMGSWTG 378

DB 366 VIEHSMARRAFPFRLAHLAALAPARTDRGDSGGGDRGGGGRVALTAPGAADAPG 425

QY 379 ---EGVHWYPOLLYRLGRLLLEEGSFHPLGMS-----GAG 410

DB 426 AGATAGIHWYSQLTQIGTWLDSBALHPLGMAVKSXSXRGAG 468

Search completed: June 5, 2000, 08:16:18
Job time: 2583 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2000, 07:55:36 ; Search time 45.88 Seconds

(without alignments)
129.332 Million cell updates/sec

Title: US-08-900-220-16

Perfect score: 2175
Sequence: 1 MSPARLRPRHLHCLVLLLL.....GRLLLEGGSFHPLGMSGAGS 411Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 145308 seqs, 14437401 residues

Total number of hits satisfying chosen parameters: 145308

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summariesDatabase : Issued Patents, AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/PTCUS.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1686	77.5	336	1	US-08-176-427B-6
2	1686	77.5	336	2	US-08-356-060A-10
3	1646	75.7	313	2	US-08-356-060A-14
4	1279.5	58.8	437	4	PCT-US95-15463-20
5	1279.5	58.8	437	4	PCT-US95-15463-20
6	1276.5	58.7	437	4	US-08-176-427B-8
7	1276.5	58.7	437	2	US-08-356-060A-11
8	1274	58.6	418	4	PCT-US95-15463-18
9	1274	58.6	418	4	PCT-US95-15463-18
10	1271	58.4	475	2	US-08-356-060A-13
11	1270.5	58.4	425	1	US-08-176-427B-2
12	1270.5	58.4	425	2	US-08-356-060A-8
13	1268.5	58.3	462	1	US-08-748-591-4
14	1268.5	58.3	462	1	US-08-748-591-9
15	1246.5	57.3	437	1	PCT-US95-02315-2
16	1227	56.4	418	1	US-08-176-427B-10
17	1227	56.4	418	2	US-08-356-060A-12
18	1226.5	56.4	416	4	PCT-US95-15463-17
19	1226.5	56.4	416	4	PCT-US95-15463-17
20	1190.5	54.7	425	4	PCT-US95-15463-19
21	1190.5	54.7	425	4	PCT-US95-15463-19
22	1175	54.0	396	1	US-08-176-427B-4
23	1175	54.0	396	2	US-08-356-060A-9
24	893.5	41.1	471	1	US-08-176-427B-11
25	893.5	41.1	471	2	US-08-356-060A-34
26	775.5	35.7	165	2	US-08-356-060A-41
27	610.5	28.1	167	2	US-08-356-060A-41
28	381	17.5	73	1	US-08-176-427B-13
29	381	17.5	73	2	US-08-356-060A-36

30	346	15.9	73	1	US-08-176-427B-12	Sequence 12, Appl
31	346	15.9	73	2	US-08-356-060A-35	Sequence 35, Appl
32	315	14.5	64	1	US-08-176-427B-15	Sequence 15, Appl
33	315	14.5	64	2	US-08-356-060A-15	Sequence 15, Appl
34	314	14.4	64	1	US-08-176-427B-17	Sequence 17, Appl
35	314	14.4	64	2	US-08-356-060A-17	Sequence 17, Appl
36	300.5	13.8	65	1	US-08-176-427B-16	Sequence 16, Appl
37	300.5	13.8	65	2	US-08-356-060A-16	Sequence 16, Appl
38	263	12.1	64	1	US-08-176-427B-14	Sequence 14, Appl
39	263	12.1	64	2	US-08-356-060A-37	Sequence 37, Appl
40	167	7.7	30	1	US-08-748-591-3	Sequence 3, Appl
41	164	7.5	30	1	US-08-748-591-2	Sequence 2, Appl
42	147	6.8	30	1	US-08-748-591-1	Sequence 1, Appl
43	145	6.7	27	1	US-08-748-591-8	Sequence 8, Appl
44	139	6.4	27	1	US-08-748-591-7	Sequence 7, Appl
45	128	5.9	27	1	US-08-748-591-6	Sequence 6, Appl

ALIGNMENTS

```
RESULT 1
US-08-176-427B-6
; Sequence 6, Application US/08176427B
; Patent No. 5789543
;
; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tablin, Clifford J.
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; TITLE OF INVENTION: Proteins and Uses Related Thereto
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/176,427B
; FILING DATE: 30-DEC-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 336 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-176-427B-6
```

Query Match 77.5% Score 1686; DB 1; Length 336;
Best Local Similarity 94.9%; Pred. No. 2; le-168;
Matches 319; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 76 ERFKELTPNYNDIFKKEENTGADRLMTORCKDLNLAISVNMQWGVKLRVTEGMD 135
DB 1 ERFKELTPNYNDIFKKEENTGADRLMTORCKDLNLAISVNMQWGVKLRVTEGMD 60
QY 136 DGHSEESIHYEGRAVDITTSDBDRNKYGLLARLAVEGFDWVYYSKRAHVHCYSKSEHS 195

Db 61 DGHSEESLHYGRADVITTSRDRNKKYGLLARLAVAGFDMVYTESKAHVCSVSEHS 120
QY 196 AAKTGCGCPAGAOVRLESGARVALSAVRPGDRLVAMGEDGSPFSDVLIFFLDREPHRLR 255
Db 121 AAKTGCGCPAGAOVRLENGERVALSAVKGDRVLAMGEDGTPFSDVLIFFLDREPHRLR 180
QY 256 AFOVITODPPRRLLATPRLHLLFTADNHTPEPAFRATFASHVQPGQYLVAGVPGLOPA 315
Db 181 AFOVITODPPRRLLATPRLHLLFTADNHTPEPAFRATFASHVQPGQYLVAGVPGLOPA 240
QY 316 RVAAVSTHVALGAYAPLTKHGTLVVEDVYVASCFAAVADHLLAQLAFWPLRFLPSLAWGSM 375
Db 241 RVAAVSTHVALGAYAPLTKHGTLVVEDVYVASCFAAVADHLLAQLAFWPLRFLPSLAWGSM 300
QY 376 TPGEVHWYPMOLLYRLGRLLLEEGSFHPLGMSGAGS 411
Db 301 TPGEVHWYPMOLLYRLGRLLLEEGSFHPLGMSGAGS 336

RESULT 2

US-08-356-060A-10
; Sequence 10, Application US/08356060A
; Patent No. 5844079
; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tablin, Clifford J.
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; TITLE OF INVENTION: Proteins and Uses Related Thereto
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,060A
; FILING DATE: 14-DEC-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,427
; FILING DATE: 30-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-006CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 336 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-356-060A-10

Query Match 77.5%; Score 1686; DB 2; Length 336;
Best Local Similarity 94.9%; Pred. No. 2.1e-168;
Matches 319; Conservative 9; Mismatches 8; Indels 0; Gaps 0;
QY 76 EREKELTPNYPDIIFKDEENTGADRLMTORCKDRNLNSLAISVANOMPVGKLVTEGMD 135
Db 1 EREKELTPNYPDIIFKDEENTGADRLMTORCKDRNLNSLAISVANOMPVGKLVTEGMD 60

QY 136 DGHSEESLHYGRADVITTSRDRNKKYGLLARLAVAGFDMVYTESKAHVCSVSEHS 195
Db 61 DGHSEESLHYGRADVITTSRDRNKKYGLLARLAVAGFDMVYTESKAHVCSVSEHS 120
QY 196 AAKTGCGCPAGAOVRLESGARVALSAVRPGDRLVAMGEDGSPFSDVLIFFLDREPHRLR 255
Db 121 AAKTGCGCPAGAOVRLENGERVALSAVKGDRVLAMGEDGTPFSDVLIFFLDREPHRLR 180
QY 256 AFOVITODPPRRLLATPRLHLLFTADNHTPEPAFRATFASHVQPGQYLVAGVPGLOPA 315
Db 181 AFOVITODPPRRLLATPRLHLLFTADNHTPEPAFRATFASHVQPGQYLVAGVPGLOPA 240
QY 316 RVAAVSTHVALGAYAPLTKHGTLVVEDVYVASCFAAVADHLLAQLAFWPLRFLPSLAWGSM 375
Db 241 RVAAVSTHVALGAYAPLTKHGTLVVEDVYVASCFAAVADHLLAQLAFWPLRFLPSLAWGSM 300
QY 376 TPGEVHWYPMOLLYRLGRLLLEEGSFHPLGMSGAGS 411
Db 301 TPGEVHWYPMOLLYRLGRLLLEEGSFHPLGMSGAGS 336

RESULT 3

US-08-356-060A-14
; Sequence 14, Application US/08356060A
; Patent No. 5844079
; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tablin, Clifford J.
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; TITLE OF INVENTION: Proteins and Uses Related Thereto
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,060A
; FILING DATE: 14-DEC-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,427
; FILING DATE: 30-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-006CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-356-060A-14

Query Match 75.7%; Score 1646; DB 2; Length 313;
Best Local Similarity 100.0%; Pred. No. 2.8e-164;
Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 101 RMTORCKDRNLNSLAISVANOMPVGKLVTEGMDGHHSESLHYGRADVITTSRDR 160

```

Db      2 RLTQKCKDNLNLSIAVNMQGVKRLVTEGDEDDHSESHYGRVADITTSRDR 61
QY      161 NKYGLLARLAVEGFDWVYVESKAHVCSVSEHSAAKTGCCPGAGQVRLSEGARVAL 220
Db      62 NKYGLLARLAVEGFDWVYVESKAHVCSVSEHSAAKTGCCPGAGQVRLSEGARVAL 121
QY      221 SAVRPGGRVLAIVEDGSPFSDVLIIFLDREPHRLRAFOVETODPPRRLLTPAHLLEFTA 280
Db      122 SAVRPGGRVLAIVEDGSPFSDVLIIFLDREPHRLRAFOVETODPPRRLLTPAHLLEFTA 181
QY      281 DNHTEPARFRATFASHVQGOVYLVAGVPGILOPARVAVSTHVALGAVAPLTKHGTLLV 340
Db      182 DNHTEPARFRATFASHVQGOVYLVAGVPGILOPARVAVSTHVALGAVAPLTKHGTLLV 241
QY      341 EDVYASCFEAAVADHHLAQLAFWPLRLPHSLAMGSWTPGEGVHWYPLLRYLGRLLLEGS 400
Db      242 EDVYASCFEAAVADHHLAQLAFWPLRLPHSLAMGSWTPGEGVHWYPLLRYLGRLLLEGS 301
QY      401 FPHLGMGAGS 411
Db      302 FPHLGMGAGS 312

```

RESULT 4
PCT-US95-15463-20
Sequence 20, Application PC/TUS9515463

```

GENERAL INFORMATION:
APPLICANT: The Johns Hopkins University School of Medicine
TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: U.S.A.
ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/15463
FILING DATE: 01-DEC-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Halle, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/080M01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-15463-20

```

Query Match 58.8%; Score 1279.5; DB 4; Length 437;
Best Local Similarity 61.1%; Pred. No. 1.2e-125;
Matches 262; Conservative 39; Mismatches 91; Indels 37; Gaps 8;

```

QY      13 CLVLL--LLVPAAMCGGRVYSGRRPRKLVPLAYQFSPNPEKTLGASGRYEG 69
Db      8 CFLVILASSLLVCPGL-ACGPGRGFG-KRRHPKRLTPLAYKQFIPNVAEKTILGASGRYEG 65
QY      70 KIASSERFKELTPNYPNDIIFKDEENTGADRLTORCKDNLNLSIAVNMQGVKRLV 129

```

```

Db      66 KIRNSERFKELTPNYPNDIIFKDEENTGADRLTORCKDNLNLSIAVNMQGVKRLV 125
QY      130 TEGWDEDGHHSESLAYEGRVADITTSRDRNKYGLIARLAVEGFDWVYVESKAHVCS 189
Db      126 TEGWDEDGHHSESLAYEGRVADITTSRDRSKYGMILARLAVEGFDWVYVESKAHVCS 185
QY      190 VKSEHSAAKTGCCPGAGQVRLSEGARVALSAVRPDRVLAMGEDSPFSDVLIIFLDR 249
Db      186 VKSEHSAAKTGCCPGAGQVRLSEGARVALSAVRPDRVLAMGEDSPFSDVLIIFLDR 245
QY      250 EPHRLRAFOVETODPPRRLLTPAHLLEFTA-DNHTEPARFRATFASHVQGOVYLVAG 308
Db      246 DEGAKKTFYIETLPERRLTLRAHLLEFTAAPHNDSCPTGSPALFASRRYRPGQRYVVA 305
QY      309 VPG---LQPARVAAS-THVALGAVAPLTKHGTLLVEDVYASCFEAAVADHHLAQLAFWP 363
Db      306 ERGDRLRLPAAVHSTLREEGAVAPLTKHGTLLINRLVASCYAVIEEHSWHRAP 365
QY      364 LRLFHSL-----AMGSWTPGEGVHWYPLLRYLGRLLLEE 398
Db      366 FRLAHALLAALAPARDGGGGGSIIPAQSAATEARGA-EPTAGIHWYSQLYHIGTWLDS 424
QY      399 GSFHPLGMS 407
Db      425 ETMHPGLMA 433

```

RESULT 5
PCT-US95-15923-20
Sequence 20, Application PC/TUS9515923

```

GENERAL INFORMATION:
APPLICANT: The Johns Hopkins University School of Medicine, et al.
TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: U.S.A.
ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/15923
FILING DATE: 04-DEC-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Halle, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/043M01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-15923-20

```

Query Match 58.8%; Score 1279.5; DB 4; Length 437;
Best Local Similarity 61.1%; Pred. No. 1.2e-125;
Matches 262; Conservative 39; Mismatches 91; Indels 37; Gaps 8;

```

QY      13 CLVLL--LLVPAAMCGGRVYSGRRPRKLVPLAYQFSPNPEKTLGASGRYEG 69
Db      8 CFLVILASSLLVCPGL-ACGPGRGFG-KRRHPKRLTPLAYKQFIPNVAEKTILGASGRYEG 65

```

[illegible]

RESULT 6
US-08-176-427B-8
; Sequence 8, Application US/08176427B
Data was 5700643

```

GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tablin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
NUMBER OF INVENTION: Proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHYE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,427B
FILING DATE: 30-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMI-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-176-427B-8

```

Query Match	58.7%;	Score 1276.5;	DB 1;	Length 437;
Best Local Similarity	-60.8%;	Pred. No. 2.4e-125;		
Matches 261;	Conservative 40;	Mismatches 91;	Indels 37;	Gaps 8;

QY	13	CLVLL-----LLVVPAAACGGGRRVVGSRRRPRKLVPLATYKOFSPNYPEKTLASGRREG	69
Db	8	CFVYLASSLLVCGL-ACSGRGRG-KRRHKPLTPLYAQOFLPNAEKTLAGSGREG	65
QY	70	KIASSERFKELTPYNDIIFKDEENTGADRLMTORCKDRLNSLAISVNMQGVKRV	129
Db	66	KITRSEFFKELTPYNDIIFKDEENTGADRLMTORCKDKLNLALISVNMQGVRLRV	125
QY	130	TEGMEDDHHEESIHYBGRVADTTSPDRNKXGLARLAVEGPMVYSESAHYHCS	189
Db	126	TEGMEDDHHEESIHYBGRVADTTSPDRNKXGLARLAVEGPMVYSESAHYHCS	185
QY	190	VKSEHSAAKKGCGCFPAQAQYRLSESGARVALSAVRPDRVLAMEDSPTFSDLIFLDR	245
Db	166	VKAENSAVAKSGCGCFPGSATVHLBQGGKLVKDLRPGDRVLADQGRLLYSDFLFLDR	241
QY	250	EPHRLRAQVLETDDPRRLALTVAHLFLTA-DNHTPEARFRATFSHYQPGCYLVAC	309
Db	246	DEGAKKVFYETLEPRBRLLTVAHLFLTAHPAHNDSDPTGPSLLEFSRVRPGGRVVA	305
QY	309	VPG-----LQPARVAAS-TVHALGATVAPLTGHTVLVEDVAVSCFAVADHHLQLAFWP	363
Db	306	ERGGRRLLPRAAVSVTLREEDAGATVAPLTGHTILINRYLAVSAYVIEESVNHRRFAP	363
QY	364	LRLFLSL-----ANGSNTPGEGVHWYDQLLYRLGRLLEE	398
Db	366	FRLLAHLAALAPARTDGGGGSIPAAQSATEARGA-EPTAGIHWYSQLLYHTIGTWLDS	424
QY	399	GSFHLIGMS 407	
Db	425	ETMHPLGMA 433	

RESULT 7
US-08-356-060A-11
; Sequence 11, Application US/08356060A
; Patent No. 5844079

APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tablin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
TITLE OF INVENTION: Proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,060A
FILING DATE: 14-DEC-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMT-006CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400

Thu Jun 8 15:52:35 2000

us-08-900-220-16.raii

Page 5

```

: TELEFAX: (617) 227-5941
: INFORMATION FOR SEQ ID NO: 11
:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 437 amino acids
:
: TYPE: amino acid
:
: TOPOLOGY: linear
:
: MOLECULE TYPE: protein
:
: US-08-356-060A-11

```

Query Match	58.7%	Score 1276.5;	DB-2	Length 437;
Best Local Similarity	60.8%	Pred. No. 2.4e-125;		
Matches 261; Conservative	40;	Mismatches 91;	Indels 37;	Gaps 8

QY	13	CLVLL--LLVVAAMCGGPRVVGSSRRPRKRLVLAKEPSPNPEKTLGASRGEG	69
Db	8	CEVLIALSSLVCGL-ACGPGRGFG-KRRPKKLTFLVAKQFLPNVAEKTLGASRGEG	65
QY	70	KIASSEFFKLTFRNYNDITFKDEENTGADRLMTQCKDRLSLAISVNMOPGVKRLV	1299
Db	66	KITRSEFFKLTFRNYNDITFKDEENTGADRLMTQCKDRLSLAISVNMOPGVKRLV	1255
QY	130	TEGMEDECHHEESLHYEGRAVDITTSDBRKNKYGLLARLAVEGFMVYSEKAVHCS	189
Db	126	TEGMEDECHHEESLHYEGRAVDITTSDBRKNKYGLLARLAVEGFMVYSEKAVHCS	1855
QY	190	VKSESAAKATGGCGFPAGAQVRLSESGARVA,LSAVRPGDRLVAMEDESPFSDVLIPLDR	249
Db	186	VKAESVAAKSGCGCFPGSATVHLBOGGKILVKDLRPGDRLVAAOQGRLLYSDPLFLDR	2455
QY	250	EPHRLAAQVLETODPPRRLLTAAHLFTA-DNHTPEARFRATFTFSHYOPQCYLVAG	308
Db	246	DEGAKKQVYLETLEPRBRLLLTAAHLFTVAPNHDSGTPGPSLFFASRRVRPGGRVYVA	305
QY	309	VPG----LQPARVAAS-THVALATVAPLTNHTGLVEDVVASCFEAAVDHHLAQLAFWP	3635
Db	306	ERGGDRRLLPAAVHSYTLREERAGATAPLTNHTGILLNRYLASCAYATIEESNAHRFAP	3655
QY	364	LRLFLSL-----AWGSWTPGEGVHWYDQLLYRLGRLLEE	398
Db	366	FLRLAALLAALAPARTDGGGGSIPAAQASATEARGA-EPTAGIMYSQLLYHITGTLDS	4244
QY	369	GSFHLIGMS 407	
Db	425	ETMHPLGMA 433	

RESULT 8
PCT-US95-15463-18
Sequence 18, Application PC/TUS9515463
GENERAL INFORMATION:
APPLICANT: The Johns Hopkins University School of Medicine
TITLE OF INVENTION: NOVEL HEDEHOG-DERIVED POLYPEPTIDES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: LA JOLLA
STATE: CA
COUNTRY: U.S.A.
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/15463
FILING DATE: 01-DEC-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347

```

? REFERENCE/DOCKET NUMBER: 07265/080001
?
? TELECOMMUNICATION INFORMATION:
?
? TELEPHONE: 619/678-5070
?
? TELEFAX: 619/678-5099
?
? INFORMATION FOR SEQ ID NO: 18:
?
? SEQUENCE CHARACTERISTICS:
?
?     LENGTH: 418 amino acids
?     TYPE: amino acid
?     TOPOLOGY: linear
?     MOLECULE TYPE: protein
?
PCT-US95-15463-18

```

Query Match	58.68	Score 1274	DB 4	Length 418
Best Local Similarity	61.28	Pred. No. 4	1e-125	
Matches 254	Conservative 46	Mismatches 101	Indels 14	Gaps 4

```

0Y 5 RLRLPHLCVLLLLLVLPAMCGCPRWVGSRRRPKRLVPLAKOFSNVVEKTIGAS 64
1 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2 RLRLVLLVSLTLSTLIV -SGLACGPGRGV -RRRHKKLPLPLAKOFLINVAEKTIGAS 59

0Y 65 GRYECKIARSSERREKELTPNINPDIIFKDEBENTGADRLMTORCKDRUNSLAISVNONPG 1244
1 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 60 GRYECKIIRNSEERREKELTPNINPDIIFKDEBENTGADRLMTORCKDKRLNSLAISVNHWP 119

0Y 125 VKLRTBEMDDGHHSESLIYEBERAVDITTSDDKRNKYGGLANLVAEAFDWWYYESKA 1044
1 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 120 VKLRTBEMDDGHHFEESLIYEBERAVDITTSDDKRYGGLSLRLVAEAFDWWYYESKA 179

0Y 185 HVHCSEKSEHSAAKTGCGPAGAQRVRESGARVALNAVRRPGDVLNMGESPTFSDVL 2444
1 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 180 HHCYVKAENSVAAKSGCGFPGSLVLSIQDGGQAVADLNPBGKVLNLAADSGNLVYFSDFI 2339

0Y 245 IFLDREPHRLAFOVIEOTDPRRLATLPAHLFTADNHTHEPAARFARATFASHYPOGOY 304
1 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 240 METDREPTTRRVEFYIETQEPVEKITLTAHLLEFVLNSTEDLHTMTAAVASYRAGOV 299

0Y 305 LVYAGVPG-LQPARPAVSTHVALAVALPLTHGCLVYEDVYASCFAAVADHHLAQLAFWP 366
1 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 300 MVYDSDGGLKSVIYORITTEBORSRFPVTAHGIYVDRIASCAVYIEDOGLAHAFAP 359

0Y 364 LRLFH-----SLMGSWTPGGGVHYPOLLVRLGLRLLEBGSFHPGLGS 407
1 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 360 ARLIYYSSSTLPONSSSRSNATLQOEBVHTSKLLIOMGWTLDSSNMLHPGLGS 414

```

RESULT 9
 PCT-US95-15923-18
 Sequence 18, Application PC/TUS9515923
 GENERAL INFORMATION:
 APPLICANT: The Johns Hopkins University School of Medicine, et al
 TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 4225 Executive Square, Suite 1400
 CITY: La Jolla
 STATE: CA
 COUNTRY: U.S.A.
 ZIP: 92037
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/15923
 FILING DATE: 04-DEC-1995
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Haile, Lisa A.
 REGISTRATION NUMBER: 38,347
 REFERENCE/DOCKET NUMBER: 07265/043WO1

```

: TELECOMMUNICATION INFORMATION
: TELEPHONE: 619/678-5070
: TELEFAX: 619/678-5099
: INFORMATION FOR SEQ ID NO: 18:
: SEQUENCE CHARACTERISTICS
: LENGTH: 418 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: PCT-US95-15923-18

```

Query Match	58.68%	Score 1274	DB 4	Length 416
Best Local Similarity	61.28%	Pred No. 4	1e-125	
Matches 254	Conservative 46	Mismatches 101	Indels 14	Gaps 4

```

QY      5 RLRPLHLFCVLVLLLVLLVPAAMGCGPGVWVSRRRPRLKVLAVLAKOFSFWPVRKLTGAS 64
D      2 RLRLRVLLVSLTSLTV -SGLACGPGGYG -RRRHPRKRLPLAVLKQFIRVAEKITGAS 59
QY      65 GRYEGRKIRASSERRERKELTPNPNPDIIFKEDEBNTGADRLMTORCKDRJNSLAISYMNWPG 1244
D      60 GRYEGRKIRNRSERERKELTPNPNPDIIFKEDEBNTGADRLMTORCKDKJNSLAISYMNWPG 119
QY      125 VKLRATEBWDGDDGHHSESLHYEGRAVDITTSDDRNKTCGLAFLAVAGEFDWYIYESKA 1844
D      120 VKLRATEBWDGDDGHHSESLHYEGRAVDITTSDDRNKTCGLAFLAVAGEFDWYIYESKA 179
QY      185 HHVCSVKSESAANKTGCGCFAGAOVRLSESGARVALSVRGDGVNLMGEGDGPTEFDVL 2444
D      180 HHCSVKAKENSVAAKSGCGCFPGSLVLSIQDGGQAAVNDLNPBGDVLVLAADSGNLTVSDFI 2299
QY      245 IFLDREPHRLAFOVIETODPPRRALATPAHLITLTADNHTPEPARFATPEASHQPOQYV 304
D      240 METDRDSTTRAVFYIEQPEVEKITTLTAHLITLVLDNSTEDTLTMAAVAASVBAQKV 299
QY      305 LVAGVPG -LQPARVAANSTHVALLAFLKNGTLVVEDVVAASCEAAVADHHLAQLAFLWP 3633
D      300 MYVVDSSGGLKSVIVORITTEORBSFPAVTHGCTIVDRILASCAVYIEQGLHLHAFAP 359
QY      364 LRLFLP-----SLAMSGWTPGGSVHHYPLLRLRLGHLLEBSFPLIGNS 407
D      360 ARLLTVYSSSLFPONSSSRSNATLQGEVWHSYSLTLVOMGVLKLDSSNMPLIGNS 414

```

RESULT 10
 US-08-356-060A-13
 Sequence 13, Application US/08356060A
 Patent No. 5844079
 GENERAL INFORMATION:
 APPLICANT: Ingham, Phillip W.
 APPLICANT: McMahon, Andrew P.
 APPLICANT: Tabin, Clifford J.
 TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
 TITLE OF INVENTION: Proteins and Uses Related Thereto
 NUMBER OF SEQUENCES: 47
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LAHIVE & COCKFIELD
 STREET: 60 State Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII(text)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/356,060A
 FILING DATE: 14-DEC-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:

```

? APPLICATION NUMBER: US 08/176,427
?
? FILING DATE: 30-DEC-1993
?
? ATTORNEY/AGENT INFORMATION:
?
?   NAME: Vincent, Matthew P.
?   REGISTRATION NUMBER: 36,709
?   REFERENCE/DOCKET NUMBER: HMI-0060
?   TELECOMMUNICATION INFORMATION:
?   TELEPHONE: (617) 221-7400
?   TELEFAX: (617) 227-5941
?   INFORMATION FOR SEQ ID NO: 13:
?     SOURCE CHARACTERISTICS:
?       LENGTH: 475 amino acids
?       TYPE: amino acid
?       TOPOLOGY: linear
?       MOLECULE TYPE: protein
?
? US-08-356-060A-13

```

Query Match	58.4%;	Score 1271;	DB 2;	Length 475;
Best Local Similarity	57.7%;	Pred. No. 1e-124;		
Matches 267;	Conservative 45;	Mismatches 85;	Indels 66;	Gaps 10;

QY	13	CLVLLLL- LVTPAAMCCGGRRVYSSRRPRPKLPLATQKQSPNNVEKTLGASGRYEG	70
Db	7	CLLVIVSSLLVCGSLACGRRGRC- KRHRPKLPLATQKQFIVNAEKTLGASGRIEGK	65
QY	71	TARSSERKELTPYNDPIIFKDEENTGADRLMTQRCCKDRLSLAISVMNQPGVKLRVT	130
Db	66	ISRRSEEFKELTPYNDPIIFKDEENTGADRLMTQRCCKDRLSLAISVMNQPGVKLRVT	125
QY	131	EGMEDGHSEESIHYEGRAVDITTSORDNKKYGLLARLVAEGFDWVYYSKRAHYICSV	190
Db	126	EGMEDGHSEESIHYEGRAVDITTSORDNKKYGLLARLVAEAFDWWYYSKRAHYICSV	185
QY	191	KSESAAAKTGGCPCPAGAOYRLSESGARVALSVAPRGDVLAMGEGDSPTSDVLIIFDRE	250
Db	186	KAENSVAAKSGCGCPGSAVTYHLEGGCTKRYKVDLSPGRVYLAADQGRLLVSDFLTFEDRD	245
QY	251	PHRLRAQVLIETQDPPRRLLATPAHLFLTA---DNHT-EPAARF-----RATE	290
Db	246	DGAKKVVYVYIETREPRERRELLTAHLFLVPAHNDSATGEEPEASSGSGPPSGALGRPALF	305
QY	295	ASHYQPGQYVVA----GVPGLOPARVAAYS-THVATGAVALPKHGTIVVEDVVASCPA	345
Db	306	ASRYRPGQRYVVAERBGDRRLRLPAVAHSVTLSEAAAGATAPLTAQSTILINRYLASCYA	365
QY	350	AVADHDLAQLAFWDLRLFHSI-----AWGSWTPG	378
Db	366	VIEHSNAHRAFAFRLLAHMLAALAPARDTRGDSGGGDRGGCGGRVALTAPRADAPG	422
QY	379	-----EGVHWYRQLLYKRLGLLLLEEGSHPLQMS-----GAG	410
Db	426	AGATAGIHWYSQLLTYIGTWLMLSEALHGRVGMAYKSSXSRGAG	468

RESULT 11
 US-08-176-427B-2
 Sequence 2, Application US/08176427B
 Patent No. 5785543
 GENERAL INFORMATION:
 APPLICANT: Ingham, Phillip W.
 APPLICANT: McMahon, Andrew P.
 APPLICANT: Tablin, Clifford J.
 TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
 TITLE OF INVENTION: Proteins and Uses Related Thereto
 NUMBER OF SEQUENCES: 33
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LAHYE & COCKFIELD
 STREET: 60 State Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,427B
FILING DATE: 30-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMT-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 425 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-176-427B-2
```

Query Match 58.4%; Score 1270.5; DB 1; Length 425;
Best Local Similarity 63.0%; Pred. No. 9.7e-125;
Matches 257; Conservative 38; Mismatches 96; Indels 17; Gaps 5;

```
QY 20 LVYPAAMGCGPGRVSGRRPRPKLVPLAYKOFSPVPEKTTGASGRYEGKTLARSSERFK 79
DB 19 LVYSSGLTGPGRGIG-KRRHFKLTPLAYKOFIPVNAEKTLGASGRYEGKTLTRNSERK 77
QY 80 ELTPNPNPIIFKDEBENTGADRLMTORCKDRNLNSLAISVMNOMPVGKLVRTGEMDEGHH 139
DB 78 ELTPNPNPIIFKDEBENTGADRLMTORCKDKLNALASVMNOMPVGKLVRTGEMDEGHH 137
QY 140 SEESLHYEGRADVITTSDDRNKYGILARLAVAGFDWVYVESKAHVCSVKSSEHSAAK 199
DB 138 SEESLHYEGRADVITTSDDRNKYGILARLAVAGFDWVYVESKAHVCSVKSSEHSAAK 197
QY 200 TGCGPAGAGVPLESGARVALSAVRPGDRLVAMGEDGSPTEFSDVLTFLDRPHRLRAFQY 259
DB 198 SGCGPFGSATVHLHGGTGLVKDLSGDRVLADADGRLYSDFTFLDRMDSRRLFTY 257
QY 260 IETODPRRLATLPAHLFTADNHTPEPA-----RFRATPASHVQPGQYLVAGVPLQ-- 314
DB 258 IETODPRRLATLPAHLFTADNHTPEPA-----RFRATPASHVQPGQYLVAGVPLQ-- 317
QY 314 PARVAAYS-THVALGAYADLTGKGLTVEDVYVASCFAVAADHLLAQLAEMPLRLFHSGLAM 372
DB 318 PASVHSVLSREESAGAYADLTGKGLTVEDVYVASCFAVAADHLLAQLAEMPLRLFHSGLAM 377
QY 373 G-----SWTPGEGVHWYPOLLYRLGRLLDEBSFPLGMSGAGS 411
DB 378 ALCPDGAIPTAATTTGTHWYSRLYRIGSWLDGDLPLGKLVADAPAS 425
```

RESULT 12
US-08-356-060A-8
Sequence 8, Application US/08356060A
Patent No. 5844079
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tablin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston

```
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,060A
FILING DATE: 14-DEC-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMT-006CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 425 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-356-060A-8
```

Query Match 58.4%; Score 1270.5; DB 2; Length 425;
Best Local Similarity 63.0%; Pred. No. 9.7e-125;
Matches 257; Conservative 38; Mismatches 96; Indels 17; Gaps 5;

```
QY 20 LVYPAAMGCGPGRVSGRRPRPKLVPLAYKOFSPVPEKTTGASGRYEGKTLARSSERFK 79
DB 19 LVYSSGLTGPGRGIG-KRRHFKLTPLAYKOFIPVNAEKTLGASGRYEGKTLTRNSERK 77
QY 80 ELTPNPNPIIFKDEBENTGADRLMTORCKDRNLNSLAISVMNOMPVGKLVRTGEMDEGHH 139
DB 78 ELTPNPNPIIFKDEBENTGADRLMTORCKDKLNALASVMNOMPVGKLVRTGEMDEGHH 137
QY 140 SEESLHYEGRADVITTSDDRNKYGILARLAVAGFDWVYVESKAHVCSVKSSEHSAAK 199
DB 138 SEESLHYEGRADVITTSDDRNKYGILARLAVAGFDWVYVESKAHVCSVKSSEHSAAK 197
QY 200 TGCGPAGAGVPLESGARVALSAVRPGDRLVAMGEDGSPTEFSDVLTFLDRPHRLRAFQY 259
DB 198 SGCGPFGSATVHLHGGTGLVKDLSGDRVLADADGRLYSDFTFLDRMDSRRLFTY 257
QY 260 IETODPRRLATLPAHLFTADNHTPEPA-----RFRATPASHVQPGQYLVAGVPLQ-- 314
DB 258 IETODPRRLATLPAHLFTADNHTPEPA-----RFRATPASHVQPGQYLVAGVPLQ-- 317
QY 314 PARVAAYS-THVALGAYADLTGKGLTVEDVYVASCFAVAADHLLAQLAEMPLRLFHSGLAM 372
DB 318 PASVHSVLSREESAGAYADLTGKGLTVEDVYVASCFAVAADHLLAQLAEMPLRLFHSGLAM 377
QY 373 G-----SWTPGEGVHWYPOLLYRLGRLLDEBSFPLGMSGAGS 411
DB 378 ALCPDGAIPTAATTTGTHWYSRLYRIGSWLDGDLPLGKLVADAPAS 425
```

RESULT 13
US-08-748-591-4
Sequence 4, Application US/08748591
Patent No. 5759811
GENERAL INFORMATION:
APPLICANT: Epstein, Eryn
APPLICANT: Hu, Zhilan
APPLICANT: Bonifas, Jeanette
TITLE OF INVENTION: Mutant Human Hedgehog Gene

QY 379 ----EGVHWYQQLYRLGRLLLESGSFHPLGMS 407
 DB 426 AGATAGIHWYSQLLYQIGTWLDSALHPLGMA 458

RESULT 15

PCT-US95-02315-2
 ; Sequence 2, Application PC/TUS9502315
 ; GENERAL INFORMATION:
 ; APPLICANT: Jessell, Thomas M.
 ; APPLICANT: Dodd, Jane
 ; APPLICANT: Roelink, Henk
 ; APPLICANT: Edlund, Thomas
 ; TITLE OF INVENTION: DNA ENCODING A VERTEBRATE HOMOLOG OF
 ; TITLE OF INVENTION: HEDGEOG, VHH-1, EXPRESSED BY THE NOTOCHORD, AND USES
 ; TITLE OF INVENTION: THEREOF
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Cooper & Dunham
 ; STREET: 1185 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/02315
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: John P. White
 ; REGISTRATION NUMBER: 28,678
 ; REFERENCE/DOCKET NUMBER: 45375-A-PCT
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 278-0400
 ; TELEFAX: (212) 391-0525
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 437 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; PCT-US95-02315-2

Query Match 57.3%; Score 1246.5; DB 4; Length 437;
 Best Local Similarity 59.9%; Pred. No. 3.3e-122;
 Matches 257; Conservative 40; Mismatches 95; Indels 37; Gaps 8;

QY 13 CLVLL--LLVPRAMCGGRVVGSRRRPRKLVPLAYKQFSNVPEKTLGASGRYEG 69
 DB 8 CFLVALASSLLVCPGL-ACGPGKFG-KRQHPKLTPLAYKQFIPVNAEKTIGASGRYEG 65
 QY 70 KIASSSEFKELTPYNYNDIIFKDEENTGADRLMTQRCCKRLNSLAISVMNOMPVKLRY 129
 DB 66 KITRNSERFKELTPYNYNDIIFKDEENTGADRLMTQRCCKRLNSLAISVMNOMPVKLRY 125
 QY 130 TEGWDEDHGHSSESLHYEGRAVDITTSDDRNKYGLLARLAVEAGFDWYVESKAHVCS 189
 DB 126 TEGWDEDHGHSSESLHYEGRAVDITTSDDRNKYGLLARLAVEAGFDWYVESKAHVCS 185
 QY 190 VKSEHSAAKGCGCPPAQOYRLGSGARVALSAVPRGDRVLAMGEDGSPFSDVLIFFDR 249
 DB 186 VKAENSVAKSDGCGPSGATVHLEGGGTRKLVKDLSPGDRVLAADQGRLLYSDFLTFDR 245
 QY 250 EPHRLRAFOVETODPPRRLALTPAHLFTA-DNHTPEARPRATFASHVPGQYVLVAG 308
 DB 246 DEGAKKYVYVETREPRRLTLTAHLFLVAPHNDSGPTPGSPPLFASRVRPGQRYVYA 305

QY 309 VPG----LQPARVAAS-THVALGAYAPLTRKGTLYVEDVYASCEFAVADHHLAQLAFWP 363
 DB 306 ERGDRRLPLPAVHSTVTLREBAAGAYAPLTADGTLINRYLASCYAVIEHSMHRAFAFAP 365
 QY 364 LRLFSL-----AMGSWTPGEGVHWYQQLYRLGRLLLEE 398
 DB 366 FRLAHALLAALAPARTDGGGGGSLPAPQSVAEARAGAPPA-GIHWYSQLLYHIGTWLDS 424
 QY 399 GSFFHPLGMS 407
 DB 425 ETLHPLGMA 433

Search completed: June 5, 2000, 08:17:15
 Job time: 1299 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2000, 07:57:20 ; Search time 22.84 Seconds

(Without alignments)
1055.027 Million cell updates/sec

Title: US-08-900-220-16
Perfect score: 2175
Sequence: 1 MSPARLRPRHFCVLLVLL.....GRLLLESGFHLGSGAGS 411

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 168808 seqs, 58629743 residues

Total number of hits satisfying chosen parameters: 168808

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1686	77.5	336	2 C49425	Indian hedgehog pr
2	1276.5	58.7	437	2 A49425	Sonic hedgehog pro
3	1274	58.6	418	2 A53193	hedgehog homolog v
4	1270.5	58.4	425	2 A49424	patterning protein
5	1251.5	57.5	444	2 S56765	morphogen xhh prec
6	1246.5	57.3	437	2 B53193	hedgehog homolog v
7	1227	56.4	415	2 A49426	Sonic hedgehog gen
8	1175	54.0	396	2 B49425	Desert hedgehog pr
9	893.5	41.1	471	2 A46400	segment polarity p
10	456	21.0	94	2 G02735	desert hedgehog -
11	174.5	8.0	615	2 T29550	hypothetical prote
12	173.5	8.0	868	2 T22281	hypothetical prote
13	170.5	7.8	1021	2 T23252	hypothetical prote
14	161	7.4	1226	2 T24045	hypothetical prote
15	154.5	7.1	1207	2 T23754	hypothetical prote
16	148.5	6.8	481	2 T27665	hypothetical prote
17	148	6.8	484	2 T34504	hypothetical prote
18	138.5	6.4	629	2 T19563	hypothetical prote
19	123.5	5.7	205	2 T26220	hypothetical prote
20	105	4.8	1997	2 T30874	virginiamylin S sy
21	102	4.7	751	2 T29357	hypothetical prote
22	101.5	4.7	898	2 S74903	DNA topoisomerase
23	99	4.6	1063	2 B70841	probable helz prot
24	98.5	4.5	316	2 T35006	beta-ketocycl-acyl
25	98	4.5	1222	2 G72614	probable reverse g
26	97	4.5	7463	2 T36248	CDA peptide synthe
27	96	4.4	8563	2 T30226	polyketide synthas
28	95.5	4.4	846	2 H70871	hypothetical prote
29	95	4.4	808	2 T04092	phospholipase D (E
30	95	4.4	1616	2 G70668	polyketide synthas

31	95	4.4	2117	2 T36180	CDA peptide synthe
32	94	4.3	482	2 S65760	dihydroliipoamide t
33	94	4.3	747	1 QRECFE	ferrichrome-iron t
34	93.5	4.3	861	2 B49847	nitrate reductase
35	92.5	4.3	127	2 T35336	Probable isomerase
36	92.5	4.3	413	2 B64983	Yela protein - Esc
37	92.5	4.3	638	2 F75547	anthranilate synth
38	92.5	4.3	736	2 T12963	subtilisin homolog
39	91.5	4.2	333	2 S75980	hypothetical prote
40	91	4.2	6260	2 T30228	polyketide synthas
41	91	4.2	10223	2 T30225	polyketide synthas
42	90.5	4.2	1102	2 JH0717	guanylate cyclase
43	90	4.1	382	2 T35709	hypothetical prote
44	90	4.1	609	1 KSA5L1	laccase (EC 1.10.3
45	89.5	4.1	1575	3 T18545	lysobactin synthe

ALIGNMENTS

RESULT 1
C49425
Indian hedgehog protein - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 06-Jan-1995
C:Accession: C49425
R:Reheland, Y.; Epstein, D.J.; St-Jacques, B.; Shen, L.; Mohler, J.; McMahon, J.A.; M
Cell 75, 1417-1430, 1993
A:Title: Sonic hedgehog, a member of a family of putative signaling molecules, is imp
A:Reference number: A49425; M01D:94094336
A:Accession: C49425
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-336 <ECCH>
A:Cross-references: GB:X76291

Query Match 77.5% ; Score 1686; DB 2; Length 336;
Best Local Similarity 94.9% ; Pred. No. 1.9e-135;
Matches 319; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY	76	ERFKEIPNPNYNDIIFDEENTGADRLMTQCKRRLSLAISVNNQMPGYKLRTEGMD	135
DB	1	ERFKEIPNPNYNDIIFDEENTGADRLMTQCKRRLSLAISVNNQMPGYKLRTEGMD	60
QY	136	DGHSEESLHYEGRAVDITTSRDRNRYGLARLAVAGEFDMVYTESKAHVCSKSEHS	195
DB	61	DGHSEESLHYEGRAVDITTSRDRNRYGLARLAVAGEFDMVYTESKAHVCSKSEHS	120
QY	196	AAATGCGCFPAGAOVRLSGARVALSAVRPGDRVLAMGEDSPFFSDVLIPLDRPHRLR	255
DB	121	AAATGCGCFPAGAOVRLSGARVALSAVRPGDRVLAMGEDSPFFSDVLIPLDRPHRLR	180
QY	256	AFQVETDPPRRLLALPFAHLFTADNHTPEPAFRATFASHVPOGVAVAGVGLQPA	315
DB	181	AFQVETDPPRRLLALPFAHLFTADNHTPEPAFRATFASHVPOGVAVAGVGLQPA	240
QY	316	RVAAVSTHVALGAYAPLTKHGTLLVEDVVASCFPAVADHLLAQAFWPLRLFLSLANGSW	375
DB	241	RVAAVSTHVALGAYAPLTKHGTLLVEDVVASCFPAVADHLLAQAFWPLRLFLSLANGSW	300
QY	376	TPGEGVHYRQQLYRLGLLLEBSFHLGSGAGS 411	
DB	301	TPGEGVHYRQQLYRLGLLLEBSFHLGSGAGS 336	

RESULT 2
A49425
Sonic hedgehog protein precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Dec-1999
C:Accession: A49425
R:Reheland, Y.; Epstein, D.J.; St-Jacques, B.; Shen, L.; Mohler, J.; McMahon, J.A.; M


```

Oy  250 EPHRLRFEVQETEDDPRLRLPAHLFLFADNHEPAPRFRFLRPSHPQCVYLVAGV 309
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  305 NLEQONFVQLHT-DGCAVLTVIPAHLVSVWQPESO--KLTFVFPADRIEKNQVLRDV 360
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy  310 P-GIOPARAASVSTWALGAYAPLTKEGTLVEDVVASCFAAVDHHLAQLAFWPLRLE 367
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  361 ETGELRPQRYVVKVGSVSKGVAPLTFEGEIVTVNSVAASGVAINVQSGLAHNGLAPMILL 420
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy  368 HSLANGSTPGE-----GVHWYPLLRLRGLLIEGSEFH 402
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  421 STLE-AWLPKAEQLHSSPKVYSSAQOQNGIHMYANALYKVDYVLPQSMRH 470
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
G02735
desert hedgehog - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 17-Jul-1998
C:Accession: G02735
R:Drummond, I.A.
submitted to the EMBL Data Library, June 1996
A:Reference number: H01643
A:Accession: G02735
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-94 <DRU>
A:Cross-references: EMBL:U059748; NID:g1401271; PID:g1401272
C:Genetics:
A:Gene: hDHH

```

Query Match	21.0%	Score 456;	DB 2;	Length 94;
Best Local Similarity	86.2%;	Pred No. 7.6e-32;		
Matches	81;	Conservative	11;	Mismatches 2;
			Indels	0;
			Gaps	0;

QY	1	IFKDEENTGADRLMTQRCCKDRIINSTAISVNNQMPGVKLTIVTEGMDDEGHSHSESLIHYEG	148
Db	1	IFKDEENSGARLILTEKREKRYNNALAIYMMWVMPGVRLRVTEGMDDEGHSHSESLIHYEG	60

```
QY 149 RAVDITTTSDRRNKKYGLLARLAVEAGFDWYYES 18
      ||:|||||
Db 61 RALDITTTSDRRNKKYGLLARLAVEAGFDWYYES 94
```

```

RESULT 11
129550
Hypothetical protein ZK377.1 - Caenorhabditis elegans
C.Species: Caenorhabditis elegans
C.Date: 15-Oct-1999 #sequence-revision 15-Oct-1999 #text-change 15-Oct-1999
C.Accession: T29550
R.Nhan, M.; Hawkins, J.
submitted to The EMBL Data Library, February 1997
A.Description: The sequence of C. elegans cosmid ZK377.
A.Reference number: 220639
A.Accession: T29550
A.Status: preliminary; translated from GB/EMBL/DDBJ
A.Molecule type: DNA
A.Residues: 1-615 <NHA>
A.Cross-references: PTDN:AA52656.1; GSPDB:GN00028; CESP:ZK377.1
A.Experimental source: strain Bristol N2; clone ZK377
C.Genetics:
A.Gene: CESP:ZK377.1
A.Map position: 10
A.Intons: 46/1; 70/3; 157/2; 176/1; 245/2; 314/3; 402/3; 489/2

```

	Query Match	8.0%	Score 174.5	DB 2:	length 615.
	Best Local Similarity	24.7%	Pred No	27.6%-07;	
	Matches	65;	Conservative	52;	Mismatches 111; Indels 35; Gaps 11.
OY	142 ESLHTEGRAVDTTSDRRD-NRYGLIARLAVLEGEDWYYES-KAHVCYSKSEHSAAAK	199			
DB	342 QGGSGSYGSGSYGPLADASQYNMYPAAQCQPAYAPAPAPAPAPAPAPAPAPAPASARCX	401			

```

0Y 200 T-----GG-----CFPAGAVATRESSARVALSVRGDRILAMGEOSPFSVILFD 248
Db 402 SPNGLFEGGCMQCFSDMEVELEDGDKM IKDKLIGDKVLSN-DEAVETISFVIELH 459
0Y 249 REPRLRAFOVLETQDPRLALTPAHLETFADNHT-----PAARPRATFASHVOPGQ 3020
Db 460 KRDEIABFNLEIFAN-CHSIKLDNHLIYSDCRTLSDLKLVAAKRVKDDCIHVTTDS 518
0Y 303 YVLVAGVGLQPARVAAVSTHVALGAPLTKHSTLYVEDVVASCFEAAVADHHLAQLAFW 362
Db 519 NVVI-----KKRYKRSKRIEIGIYSPSTGDIIVNRLASCHSNLAKLSLOQ-TFF 570
0Y 363 PL-----RLFSLAMGSWTPEG 380
Db 571 SLYKRTSVFHNLMFPKSSTEG 593

```

RESULT 12
T22281
hypothetical protein F46B3.5 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22281
R:A:ncough, R.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19541
A:Accession: T22281
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-868 <MW>
A:Cross-references: EMBL:Z81540; PIDN:CB04405.1; GSPDB:GN00023; CESP:F46B3.5
C:Experimental source: clone F46B3
C:Genetics:
A:Gene: CESP:F46B3.5
A:Map position: 5
A:Introns: 45/3; 200/1; 235/3; 436/3; 561/3; 581/3; 696/3; 742/1; 826/3; 846/3/3

Query Match	8.0%	Score 173.5	DB 2	Length 868
Best Local Similarity	26.3%	Pred. No. 1.5e-06		
Matches	55	Conservative	33	Mismatches 90
				Indels 31
				Gaps 6

[illegible]

```

RESULT      13
T23252
hypothetical protein K02E2.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T23252
R:|loyd, C.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19716
A:Accession: T23252
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1021 <WIL>

```

A:Cross-references: EMBL:Z81560; PIDN:CA804547.1; GSPDB:GN00023; CESP:K02E2.2
A:Experimental source: clone K02E2
C:Genetics:
A:Gene: CESP:K02E2.2
A:Map position: 5
A:Introns: 41/1; 79/3; 131/3; 197/1; 262/1; 297/3; 566/3; 610/3; 735/3; 775/3; 834/3; 88

Query Match	7.8%	Score 170.5	DB 2	Length 1021
Best Local Similarity	24.3%	Pred. No. 3.3e-06		
Matches 55	Conservative 42	Mismatches 88	Indels 41	Gaps 7

```

QY      203  CFPAGACVRLTSGKRVANLSPAPCDRYLAMEODSPFSESVLLFLEDEPRRLAFOVIE 262
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      802  CFSRDTWTWTTSSGKK-RNDELEIDYVLT-ADKTLAFSAITLMIHREPVTQVDELEIK 859
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      263  QDPERRLALPAHLLEFADNHTEDA-----ARFR-----ATFASHVOP 301
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      860  -DNCKTQLDTAGHFITYATECHRYLPSKNSLSLNTSPERYRHIDTLPDSETKLASQKIG 918
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      302  QYVLVACVPGIOPARVAANVSTHVALGAYAPLTFKGTGLYVEDVYASCPRAAYADHHLAOLAF 361
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      919  ECLLIHNGDOFRMOKIDISIKSTVSTGTPSLTEGRILVANDVYASCSYEQOQNVLTTFE 978
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      362  WPLRLHSLANGSWTPGCGVHWYQOLLRYGRLLLEEGSFHPIGMS 407
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      979  WAFDRRLNVL-----LVOYREDLTLDEIEL-PIGTS 1007
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 14
T24045
hypothetical protein R08B4.1 - *Caenorhabditis elegans*
C:\Protein_Caenorhabditis\T24045.faa

QY 198 AKTGCGCPAGAGVYVELEGAGVAAVSAVPGGVYAMCEDSPFSEVDLFEDEPRRLAAV 257
 Db 1016 AATGACFSLDPTWVTPYTKK-RMDQIDIDGYVLTADLE-KTYFTPTITIMHREPEKVEEF 1073
 QY 258 QVETGDDPRRLALTPAHLLF---TADNHPEPAAR---RATFESHQPGGVYVYAAVP 310
 Db 1074 LTIVT-EGKRLRLTSHPMYRNKCKGSYQYIKMLPHGCEALFASDDEVGDCVYVLYKG 1132
 QY 311 GLQPARVAANSTHVALGAYDLTKHGTLVEDVYVASCFAAVADHHLAQLAEW 362
 Db 1133 KYRQOKLETTIRSVRGVSYPLPTNNNGIIVNDLACSYSEIDONTLQTTFFW 1184

RESULT 15
T23754
hypothetical protein T05C12.10 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_rev:15on 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T23754; T24513

R:Thomas, K. Submitted to the EMBL Data Library June 1995
A:Reference number: Z19793
A:Accession: M33754
A:Status: preliminary; translated from GR/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-1207 <MIL>

Query Match	7.1%;	Score 154.5;	DB 2;	Length 1207;
Best Local Similarity	28.2%;	Pred. NO. 9.5e-05;		
Matches	51;	Conservative	30;	Mismatches 83;
			Indels	17;
			Gaps	5;

```

OY      196  AAATGG----CFAGAOVLESGARALVASVPGGRVLAMGEDSPFSDVLIPLDRP 251
          |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
Db      950  AAGAAGGGRSCFSADSLVTTVTGQR-RDDELQIGDIYLVPSGSNVLYKXEYEMFYHREP 1008
          |||   |||   |||   |||   |||   |||   |||   |||   |||   |||

OY      252  HRLRAFOVIETODEPRRLATTPAHL-----FTADNHTPEAPAREPAFASHVOFG 301
          |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
Db      1009 KTRTFEIVLTKS-GRLLSLTGRHLLPVACESQVEDYITNNPBGIDIVAMESSYLAKRKG 1067
          |||   |||   |||   |||   |||   |||   |||   |||   |||   |||

OY      302  QYTLVAVYPGLQPA-RVAASVTHVALGANAPLTCKGTLYVEDVDVASCFEAADVADHHLAOLA 360
          :||   :||   :||   :||   :||   :||   :||   :||   :||   :||
Db      1068 ECVSLIDESGEVIADEIVRVGRMTNGIYSPMTEVSLLIDGVLSCEFHLESHPAHKLI 1127
          :||   :||   :||   :||   :||   :||   :||   :||   :||   :||

OY      361  F 361
          |
Db      1128 F 1128

```

Search completed: June 5, 2000, 08:17:52
Job time: 1232 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2000, 08:17:16 ; Search time 14.43 Seconds

(Without alignments)
867.429 Million cell updates/sec

Title: us-08-900-220-16

Perfect score: 2175
Sequence: 1 MSPARLRPLHFLVILLLL.....GRLLLEGSFHPLGMSGAGS 411

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 83857 seqs, 30454973 residues

Total number of hits satisfying chosen parameters: 83857

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt_38:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	2162	99.4	411 1	IHH_HUMAN
2	2069	95.1	411 1	IHH_MOUSE
3	1645.5	75.7	408 1	IHH_CHICK
4	1483	68.2	409 1	IHH_XENLA
5	1279.3	58.8	437 1	SHH_MOUSE
6	1274	58.6	418 1	SHH_BRARE
7	1270.5	58.4	425 1	SHH_CHICK
8	1268.5	58.3	462 1	SHH_HUMAN
9	1263	58.1	432 1	SHH_CYNPY
10	1251.5	57.5	444 1	SHH_XENLA
11	1246.5	57.3	437 1	SHH_RAT
12	1226.5	56.4	416 1	TWHH_BRARE
13	1196	55.0	412 1	IHH_BRARE
14	1175	54.0	396 1	DHH_MOUSE
15	1171.5	53.9	396 1	DHH_HUMAN
16	1117	51.4	398 1	DHH2_XENLA
17	1114.5	51.2	396 1	DHH1_XENLA
18	882.5	40.6	471 1	HH_DROME
19	870	40.0	481 1	HH_DROHY
20	555	25.5	121 1	SHH_CARAY
21	555	25.5	121 1	SHH_RASBL
22	553	25.4	121 1	SHH_RASPA
23	552	25.4	121 1	SHH_RASBL
24	552	25.4	121 1	SHH_RASBL
25	551	25.3	121 1	SHH_RASPA
26	550	25.3	121 1	SHH_DANAE
27	550	25.3	121 1	SHH_DANAE
28	550	25.3	121 1	SHH_DANAT
29	550	25.3	121 1	SHH_DANER
30	550	25.3	121 1	SHH_DANKE
31	550	25.3	121 1	SHH_DANPU
32	550	25.3	121 1	SHH_DEVDE
33	550	25.3	121 1	SHH_DEVMA
34	550	25.3	121 1	SHH_DEVPA

35	548	25.2	121 1	SHH_RASHE	P79864 rasbora het
36	546	25.1	121 1	SHH_PUNCO	P79838 puntius con
37	295	13.6	58 1	IHH_CARAY	P79693 carassius a
38	295	13.6	58 1	IHH_DANAT	O13240 danio aff.
39	295	13.6	58 1	IHH_DANKE	P79711 danio kerri
40	295	13.6	58 1	IHH_DANPU	P79719 danio pulch
41	295	13.6	58 1	IHH_DEVDE	O13243 devario dev
42	295	13.6	58 1	IHH_PUNTE	P79852 puntius tet
43	295	13.6	58 1	IHH_RASEL	P79860 rasbora ele
44	290	13.3	58 1	DHH_BRARE	P79729 brachydanio
45	290	13.3	58 1	DHH_DANKE	P79712 danio kerri

ALIGNMENTS

RESULT 1
IHH_HUMAN STANDARD; PRT; 411 AA.
ID IHH_HUMAN
AC O14623; O43322;
DT 15-JUL-1999 (Rel. 38, Created)
DR 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE INDIAN HEDGEHOG PROTEIN PRECURSOR (IHH) (HNG-2).
GN IHH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
CC Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA Tate G., Endo Y., Mitsuura T.;
RT "Human Indian Hedgehog."
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 100-411 FROM N.A.
RC TISSUE-FETAL LUNG;
RX MEDLINE; 96070431.
RA Marigo V., Roberts D.J., Lee S.M.K., Tsukurov O., Levi T.,
RA Gaster J.M., Epstein D.J., Gilbert D.J., Copeland N.G., Seidman C.E.,
RA Jenkins N.A., Seidman J.G., McMahon A.P., Tabin C.;
RT "Cloning, expression, and chromosomal location of SHH and IHH: two
human homologues of the Drosophila segment polarity gene hedgehog."
RL Genomics 28:44-51(1995).
RN [3]
RP SEQUENCE OF 124-172 FROM N.A.
RX MEDLINE; 95236997.
RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,
RA Zhao R., Seidman J.G., Fallon J.F., Beachy P.A.;
RT "Products, genetic linkage and limb patterning activity of a murine
hedgehog gene."
RL Development 120:3339-3353(1994).
CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
PATTERNING EVENTS DURING DEVELOPMENT. BINDS TO THE PATCHED (PTC)
RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHED
ENDONUCLEOTID OXIFICATION. MAY REGULATE THE BALANCE BETWEEN GROWTH
AND OSSIFICATION OF THE DEVELOPING BONES. INDUCES THE EXPRESSION
OF PARATHYROID HORMONE-RELATED PROTEIN (PTHrP) (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN EMBRYONIC LUNG, AND IN ADULT
KIDNEY AND LIVER.
CC -1- PM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY
AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U85610; AAB49692.1; ALT_INIT.
DR EMBL: X76291; CA53923.1; -.
DR MGD; MG1:96533; IHH.
DR PFAM; PF01079; Hint; 1.
DR PFAM; PF01085; HR_signal; 1.
DR PRINTS; PR00632; SONICHOG.
KW Developmental protein; Autocatalytic cleavage; Hydrolyase; Protease;
KW Signal.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 411 INDIAN HEDGEHOG PROTEIN.
FT CHAIN 28 202 INDIAN HEDGEHOG PROTEIN N-PRODUCT.
FT CHAIN 203 411 INDIAN HEDGEHOG PROTEIN C-PRODUCT.
FT SITE 202 203 CLEAVAGE (AUTO-).
FT SITE 248 248 INVOLVED IN CHOLESTEROL TRANSFER (BY
FT SITE 248 248 INVOLVED IN CHOLESTEROL TRANSFER (BY
FT SITE 272 272 INVOLVED IN AUTO-CLEAVAGE (BY
FT ACT_SITE 275 275 ESSENTIAL FOR AUTO-CLEAVAGE (BY
FT BINDING 202 202 SIMILARITY).
FT CONFLICT 383 383 CHOLESTEROL (BY SIMILARITY).
FT SEQUENCE 411 AA; 45485 MW; 08BE7AD8507C0D9B CRC64;

Query Match 95.1%; Score 2069; DB 1; Length 411;
Best Local Similarity 95.1%; Pred. No. 2.1e-166;
Matches 391; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

OY 1 MSPALRPRLHFCVLLVLPVPAAMGCGPRVSGRRPRKLYPLAKOPSPVPEKT 60
DB 1 MSPALRPRLHFCVLLVLPVPAAMGCGPRVSGRRPRKLYPLAKOPSPVPEKT 60
OY 61 LGASRGYEGKTIARSSERKELTPNPNPDIFKDEENTGADRLMTCRCKRLNSLAISVAN 120
DB 61 LGASRGYEGKTIARSSERKELTPNPNPDIFKDEENTGADRLMTCRCKRLNSLAISVAN 120
OY 121 QMPGVKLTVEGMDDEGHSESLHYEGRAVDITSDRNRKYGILLARLAVAGFDWY 180
DB 121 QMPGVKLTVEGMDDEGHSESLHYEGRAVDITSDRNRKYGILLARLAVAGFDWY 180
OY 161 ESKAVHGSVKSSEHSAAKTGCGCFAGAGVRLSGARVALSVRGDRVLAMGEDEGSPTE 240
DB 161 ESKAVHGSVKSSEHSAAKTGCGCFAGAGVRLSGARVALSVRGDRVLAMGEDEGSPTE 240
OY 241 SDVLLFLDREPRRLRAFOVETODPPRRRLATPAHLFTADNHTPEPARFRATFASHVOP 300
DB 241 SDVLLFLDREPRRLRAFOVETODPPRRRLATPAHLFTADNHTPEPARFRATFASHVOP 300
OY 301 GQYVLAVGVPGLQPARVAAVSTHVALGAYAPLTKGTLVVEDEVASCFPAVADHHLAOLA 360
DB 301 GQYVLAVGVPGLQPARVAAVSTHVALGAYAPLTKGTLVVEDEVASCFPAVADHHLAOLA 360
OY 361 FWPPLLFHSLANGSWTPGSGVHWYFQOLYRLGRLLIEESFPLGSGAGS 411
DB 361 FWPPLLFHSLANGSWTPGSGVHWYFQOLYRLGRLLIEESFPLGSGAGS 411

RESULT 3
IHH_CHICK 3
ID IHH_CHICK STANDARD; PRT; 408 AA.
AC Q98938;
DI 15-JUL-1999 (Rel. 38, Created)
DI 15-JUL-1999 (Rel. 38, Last sequence update)
DI 15-JUL-1999 (Rel. 38, Last annotation update)
DE INDIAN HEDGEHOG PROTEIN PRECURSOR (IHH).
GN IHH.
OS Gallus gallus (Chicken).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
CC Negundinae; Galliformes; Phasianidae; Gallus.
CC [11]
CC SEQUENCE FROM N.A.
CC MEDLINE; 96325423.
CC Vorkamp A., Lee K., Lanske B., Segre G.V., Kronenberg H.M.,
CC Tablin C.J.;
CC "Regulation of rate of cartilage differentiation by Indian hedgehog
CC and PTH-related protein.";
CC Science 273:613-622(1996).
CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO
CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IMPLICATED IN
CC ENDOCHONDRAL OSSIFICATION. MAY REGULATE THE BALANCE BETWEEN GROWTH
CC AND OSSIFICATION OF THE DEVELOPING BONES. INDUCES THE EXPRESSION
CC OF PARATHYROID HORMONE-RELATED PROTEIN (PTHrP).
CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN DEVELOPING MIDGUT, LUNG AND
CC CARTILAGE OF DEVELOPING LONG BONES IN THE LIMB.
CC -1- PTH: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY
CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U85611; AAC6010.1; -.
DR PFAM; PF01079; Hint; 1.
DR PFAM; PF01085; HR_signal; 1.
DR PRINTS; PR00632; SONICHOG.
KW Developmental protein; Autocatalytic cleavage; Hydrolyase; Protease;
KW Signal.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 408 INDIAN HEDGEHOG PROTEIN.
FT CHAIN 24 198 INDIAN HEDGEHOG PROTEIN N-PRODUCT.
FT CHAIN 199 408 INDIAN HEDGEHOG PROTEIN C-PRODUCT.
FT SITE 198 199 CLEAVAGE (AUTO-) (BY SIMILARITY).
FT SITE 244 244 INVOLVED IN CHOLESTEROL TRANSFER (BY
FT SITE 244 244 INVOLVED IN CHOLESTEROL TRANSFER (BY
FT SITE 268 268 INVOLVED IN AUTO-CLEAVAGE (BY
FT ACT_SITE 271 271 ESSENTIAL FOR AUTO-CLEAVAGE (BY
FT BINDING 198 198 CHOLESTEROL (BY SIMILARITY).
FT SEQUENCE 408 AA; 44829 MW; BA397AE2A9357A24 CRC64;

Query Match 75.7%; Score 1645.5; DB 1; Length 408;
Best Local Similarity 77.7%; Pred. No. 7.8e-131;
Matches 320; Conservative 21; Mismatches 66; Indels 5; Gaps 2;

OY 1 MSPALRPRLHFCVLLVLPVPAAMGCGPRVSGRRPRKLYPLAKOPSPVPEKT 60
DB 1 MSPALRPRLHFCVLLVLPVPAAMGCGPRVSGRRPRKLYPLAKOPSPVPEKT 60
OY 61 LGASRGYEGKTIARSSERKELTPNPNPDIFKDEENTGADRLMTCRCKRLNSLAISVAN 120

```

Db 57 LGASGRYEKIAANSERFKELTPNYPNDITFKDENTGADRLMTQCKRLNSLAISVNM 116
QY 121 QMPGVKIRVTEGWDEDOGHSESLHREGRAVDITTSRDNRNKKYGLLARLAVEGFMWYX 180
Db 117 QMPGVKIRVTEGWDEDOGHSESLHREGRAVDITTSRDNRNKKYGLLARLAVEGFMWYX 176
QY 181 ESKAHNCSVKSSESAAKAGCGEPAGAOVRLSSGAVALSAPRGORVLA MEDSGPTF 240
Db 177 ESKAHNCSVKSSESAAKAGCGEPAGALTEGATPTMALPGORVLA MEDGAGPTV 236
QY 241 SDVLEIFDRPHRLRAFOVETODPPRLATPRAHLFTADNHTPEAPRRATFASHVCP 300
Db 237 SDVLEIFDRPHRLRAFAVETROPRLALTPHLLFVADNMSAPRAOGRPTFASHVCP 296
QY 301 GQVLYVA-GVPGLOPARVAASVTHVALGAYAPLTKHGTLLVEDVVASCFPAVADHHLAQL 359
Db 297 GHVVLVAVGSGGLQPAEVVGVRRGTGVGAYAPLTKHGTLLVEDVVASCFPAVADHHLAQL 356
QY 360 AFMPRLRHFSLANGSWTPRGVHWYPOLYRLGLRLLEESFPLGMSGAGS 411
Db 357 AFMPRLRHFSLANGSWTPRGVHWYSGLLYRLGLRLLEESFPLGMPRAES 408

RESULT 4
ID IHH_XENLA STANDARD: PRT: 409 AA.
AC 091612;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE INDIAN HEDGEHOG PROTEIN PRECURSOR (IHH) (BAND HEDGEHOG PROTEIN) (X-
  BHH)
GN IHH OR BHH.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
  Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae;
  OC Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-EMBRYO;
RX MEDLINE: 95401852.
RA Ekker S.C., McGrew L.L., Lai C.-J., Lee J.J., von Kessler D.P.,
  RA Moon R.T., Beachy P.A.;
RT "Distinct expression and shared activities of members of the hedgehog
  RT gene family of Xenopus laevis."
RL Development 121:2337-2347(1995).
CC -1- FUNCTION: SIGNAL INVOLVED IN THE EARLY INDUCTION AND PATTERNING OF
  CC ANTERODORSAL ECTODERM, NERVOUS SYSTEM AND SOMITES. INDUCES ECTOPIC
  CC CEMENT GLAND FORMATION IN EMBRYOS.
CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
  CC CELL SURFACE. HEDGEHOG PROTEIN IS ALSO SECRETED IN EITHER CLEAVED
  CC OR UNCLEAVED FORM TO MEDIATE SIGNALING TO OTHER CELLS (BY
  CC SIMILARITY).
CC -1- DEVELOPMENTAL STAGE: DETECTABLE WITHIN THE EARLY GASTRULA AT
  CC STAGE 14 NEURULA, HIGH EXPRESSION IN NEUROECTODERM, EXPRESSED
  CC THROUGHOUT THE NEURAL PLATE AND SUBSEQUENTLY IN BOTH THE NERVOUS
  CC SYSTEM AND IN THE DERMATOME OF SOMITES.
CC -1- INDUCTION: BY ACTIVIN.
CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
  CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
  CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
  CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
  CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
  CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
  CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
  CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
  CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U26404; AAA85165.1;
DR PFAM: PF01079; Hnt; 1.
DR PRINTS: PR00632; SONCHHOG.
KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
  KW signal.
FT SIGNAL 1 23
FT CHAIN 24 409
FT CHAIN 24 197
FT CHAIN 198 409
FT SITE 197 198
FT SITE 267 267
FT ACT_SITE 270 270
FT BINDING 197 197
FT DOMAIN 32 35
SQ SEQUENCE 409 AA; 45591 MW; 6FB265367FB98627 CRC64;
  FT BINDING 197 197
  FT DOMAIN 32 35
  SQ SEQUENCE 409 AA; 45591 MW; 6FB265367FB98627 CRC64;

Query Match 68.2%; Score 1483; DB 1; Length 409;
Best Local Similarity 70.9%; Pred. No. 3.5e-11;
Matches 285; Conservative 39; Mismatches 74; Indels 4; Gaps 3;

QY 8 PRHFEVLVLLLVYPAAGCGPGRGVSGRRPRKLYPLAKOPSPNVEKTLGASGRY 67
Db 4 PKVYLLCAALLLSGAVGCGPGRVGRRR-PTKLSLSIKQSPNVEKTLGASGRY 62
QY 68 ESKIAANSERFKELTPNYPNDITFKDENTGADRLMTQCKRLNSLAISVNMQPGVKL 127
Db 63 EGRISNSRFRKELTPNYPNDITFKDEITGADRLMTQCKRLNSLAISVNMQPGVKL 122
QY 128 RTEGWDEDOGHSESLHREGRAVDITTSRDNRNKKYGLLARLAVEGFMWYTESKAHV 187
Db 123 RTEGWDEDOGHSESLHREGRAVDITTSRDNRNKKYGLLARLAVEGFMWYTESKAHV 182
QY 188 CSVKSEHSAAGTGGCFPGAGOVRLSSGAVALSAPRGORVLA MEDSGPTF 247
Db 183 CSVKSEHSAAGTGGCFPGALATLESGEKIPVQSPLRLANDNSCRPTYSPLSTL 242
QY 248 DREPHRLRAFOVETODPPRLATPRAHLFTADNHTPEAPRRATFASHVCPGYLYVA 307
Db 243 DHPREHMFQYIKQDDPHRLFTPAHLIFVSDNYSFASFOAVFASVAPGQYLYVS 302
QY 308 GYPGLOPARVAASVTHVALGAYAPLTKHGTLLVEDVVASCFPAVADHHLAQLAFPLRF 367
Db 303 NVVGLIPARVRSVNTQNTGAYAPLTKHGTLLVEDVVASCFPAVADHHLAQLAFPLRF 362
QY 368 HSLANGSWT-PGE--GVHWYPOLYRLGLRLLEESFPLGMSGAGS 406
Db 363 YNDGIATGTPSQOMCIHWYSKALYHLGLRLHGNEDFPLGI 404

RESULT 5
SHH_MOUSE
ID SHH_MOUSE STANDARD: PRT: 437 AA.
AC 062226;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE SONIC HEDGEHOG PROTEIN PRECURSOR (SHH) (HHG-1).
GN SHH OR HHG1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
  OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-C57BL/6J; TISSUE-EMBRYO;

```


RX MEDLINE: 94094334.
 RA Echehard Y., Epstein D.J., St Jacques B., Shen L., Mohler J.,
 RA McMahon J.A., McMahon A.P.;
 RT "Sonic hedgehog, a member of a family of putative signaling
 RT molecules, is implicated in the regulation of CNS polarity";
 RL Cell 75:1417-1430(1993).
 RN [2]
 RC REVISION TO 122.
 RP STRAIN-C57BL/6J;
 RA McMahon A.P.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
 RX MEDLINE: 95236997.
 RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,
 RA Zhao R., Seldin M.F., Fallon J.F., Beachy P.A.;
 RT "Products, genetic linkage and limb patterning activity of a murine
 RT hedgehog gene";
 RL Development 120:3339-3353(1994).
 RN [4]
 RP FUNCTION, AND AUTOPROTEOLYTIC CLEAVAGE.
 RX MEDLINE: 95254654.
 RA Roelink H., Porter J.A., Chiang C., Tanabe Y., Chang D.T.,
 RA Beachy P.A., Jessell T.M.;
 RT "Floor plate and motor neuron induction by different concentrations of
 RT the amino-terminal cleavage product of sonic hedgehog
 RT autoproteolysis";
 RL Cell 81:445-455(1995).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 34-195.
 RX MEDLINE: 96069744.
 RA Hall T.M.T., Porter J.A., Beachy P.A., Leahy D.J.;
 RT "A potential catalytic site revealed by the 1.7-A crystal structure of
 RT the amino-terminal signalling domain of Sonic hedgehog";
 RL Nature 378:212-216(1995).
 CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN
 CC ASSOCIATION WITH SMOOTHED (SMO), TO ACTIVATE THE TRANSCRIPTION
 CC OF TARGET GENES. ON THE ABSENCE OF SHH, PTC REPRESENTS THE
 CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER
 CC TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A
 CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED
 CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE
 CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE
 CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH
 CC FLOOR PLATE- AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD
 CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS
 CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN A NUMBER OF EMBRYONIC TISSUES
 CC INCLUDING THE NOTOCHORD, VENTRAL NEURAL TUBE, FLOOR PLATE, LUNG
 CC BUD, ZONE OF POLARIZING ACTIVITY AND POSTERIOR DISTAL MESENCHYME
 CC OF LIMBS. IN THE ADULT, EXPRESSED IN LUNG AND NEURAL RETINA.
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE DURING GASTRULATION.
 CC -1- INDUCTION: BY RETINOIC ACID.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X76290; CAA53922.1; -
 DR PDB: 1VHH; 29-JAN-96.
 DR MGD; MGI:96297; SHH.
 DR DR PFAM; PF01079; SHH; Hint; 1.
 DR DR PFAM; PF01085; HH-signal; 1.
 DR DR PRINTS; PR00632; SONICHHOG.
 KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 KW Signal; 3D-structure.
 FT SIGNAL 1 24
 FT CHAIN 25 437
 FT CHAIN 25 198
 FT CHAIN 199 437
 FT SITE 198 199
 FT SITE 244 244
 FT SITE 244 244
 FT SITE 268 268
 FT ACT_SITE 271 271
 FT BINDING 198 198
 FT DOMAIN 383 387
 SQ SEQUENCE 437 AA; 4773 MW; D0EB72F08E7860EF CRC64;
 Query Match 58.8%; Score 1279.5; DB 1; Length 437;
 Best Local Similarity 61.1%; Pred. No. 4,6e-100;
 Matches 262; Conservative 39; Mismatches 91; Indels 37; Gaps 8;
 QY 13 CLVLL--LLVPAWCGPGRVSGRRPRKLYLAYKQSPNYEKTLAGSGHYEG 69
 DB 8 CLVLLASLLVCGPL-ACGPGRGFG-KRHHPKPLPLAKQGLPNNAEKTLAGSGHYEG 65
 QY 70 KIASSEFRELPPNYPDIIFFDEENTGADRLMTCRCKRLNSLAISVNMQGVKLV 129
 DB 66 KITNSRFELPPNYPDIIFFDEENTGADRLMTCRCKRLNSLAISVNMQGVKLV 125
 QY 130 TEGDEGHHSESLAHEGRAVDITSDRNRNKYGLARLAVENGFWMYVESAHVCS 189
 DB 126 TEGDEGHHSESLAHEGRAVDITSDRNRNKYGLARLAVENGFWMYVESAHVCS 185
 QY 190 VKSEHSAATGCGFPAGAVRLESGARVALSAVRPDRVLAMGEDSPFSDVLPDR 249
 DB 186 VKAENSVAAKSGCGFPASVAYHLEGGSTKLVKDLRPEDRLADDDGRLXSDFTLDR 245
 QY 250 EPHRLAFQVIEFDPPRLALPPHLLFTA-DNTEPAARFRTAFSHVQGVYLVAG 308
 DB 246 DECAKRYFYVETLEPERELITFAHLLFVAHPNDSGPTGSPSLFASVRPQGVYVVA 305
 QY 309 VPG----LQPARVAAS-TVALGAVAPLKHGLVYEDVAVSFAVAADHHLAQLFWP 363
 DB 306 ERGGDRLLPAAVSHVTLREENGAVAPLTAHGILLNRLVASVAVIEHSMHAFAP 365
 QY 364 LRLHSH--AMGSWTPGEGVMYVQLLYRLGLLEE 398
 DB 366 FRLHALALAPARTGGGGGSPAAQASATARGA-EPFAGTHWYSQLLYHISTWILDS 424
 QY 399 GSFFHLGMS 407
 DB 425 ETMPLGMA 433
 RESULT 6
 SHH_BRARE STANDARD; PRT; 418 AA.
 ID SHH_BRARE 013170; 013171;
 AC 092008; 013170; 013171;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE SONIC HEDGEHOG PROTEIN PRECURSOR (SHH) (VHH-1).

CN SHH OR VHH1.
 OS Brachydanio rerio (zebrafish) (Zebrafish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
 CC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
 CC Cyprinidae; Cyprinidae; Rasbora; Danio.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-EMBRYO.
 RX MEDLINE: 94170375.
 RA Rosell H., Augsburger A., Heemskerk J., Korzh V., Norlin S.,
 RA Ruiz I Altaba A., Tanabe Y., Placzek M., Edlund T., Jessell T.M.,
 RA Dodd J.;
 RT "Floor plate and motor neuron induction by vhh-1, a vertebrate
 RT homolog of hedgehog expressed by the notochord.";
 RL Cell 76:761-775(1994).
 RN [2]
 RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
 RX MEDLINE: 96014264.
 RA Ekker S.C., Unger A.R., Greenstein P., von Kessler D.P., Porter J.A.,
 RA Moon R.T., Beachy P.A.;
 RT "Patterning activities of vertebrate hedgehog proteins in the
 RT developing eye and brain.";
 RL Curr. Biol. 5:944-955(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 96083328.
 RA Fietz M.J., Concordet J.-P., Barbosa R., Johnson R., Krauss S.,
 RA McMahon A.P., Tabin C., Ingham P.W.;
 RT "The hedgehog gene family in Drosophila and vertebrate development.";
 RL Development Suppl. 43-51(1994).
 RN [4]
 RP SEQUENCE OF 30-92 AND 113-170 FROM N.A.
 RC TISSUE-MUSCLE.
 RX MEDLINE: 97075114.
 RA Zardoya R., Abouheif E., Meyer A.;
 RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
 RT closely related to the zebrafish.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE
 CC NOTOCHORD THAT INDUCES SOMITE PATTERNING, DORSO-VENTRAL PATTERNING
 CC OF THE BRAIN AND EARLY PATTERNING OF THE DEVELOPING EYES. DISPLAYS
 CC FLOOR PLATE-INDUCING ACTIVITY. BINDS TO THE PATCHED (PTC) TO
 CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO
 CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH,
 CC PTC REPRESENTS THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE VENTRAL MIDLINE OF THE NEURAL
 CC TUBE AND BRAIN. ALSO FOUND IN THE NOTOCHORD AND IN DEVELOPING FIN
 CC BUD. IN THE DEVELOPING BRAIN, EXPRESSION OCCURS IN DOMAINS THAT
 CC INCLUDE A DISCRETE REGION IN THE FLOOR OF THE Diencephalon.
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE IN THE INNER CELL LAYER OF
 CC THE EMBRYONIC SHIELD DURING GASTRULATION. BY 9.5 HRS OF
 CC DEVELOPMENT, EXPRESSED IN A CONTINUOUS BAND THAT EXTENDS FROM THE
 CC TAIL TO THE HEAD. THE ANTERIOR BOUNDARY OF EXPRESSION BEING
 CC POSITIONED IN THE CENTER OF THE ANIMAL POLE ANTERIOR TO THE
 CC PRESUMPTIVE MIDBRAIN.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L27585; AAA20998.1; -
 CC DR EMBL: U30711; AAC59742.1; -
 CC DR EMBL: Z35669; CA84738.1; -
 CC DR EMBL: U51351; AAB38575.1; -
 CC DR EMBL: U51370; AAB38593.1; -
 CC DR ZFIN: ZDB-GENE-980526-166; SHH.
 CC DR PFM: PF01079; Hint: 1.
 CC DR PFM: PF01085; HH_signal; 1.
 CC DR PRINTS: PR00632; SONICHHOG.
 CC KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 CC Signal.
 CC FT SIGNAL 1 23 POTENTIAL.
 CC FT CHAIN 24 418 SONIC HEDGEHOG PROTEIN.
 CC FT CHAIN 24 197 SONIC HEDGEHOG PROTEIN N-PRODUCT.
 CC FT CHAIN 198 418 SONIC HEDGEHOG PROTEIN C-PRODUCT.
 CC FT SITE 197 198 CLEAVAGE (AUTO-).
 CC FT SITE 243 243 INVOLVED IN CHOLESTEROL TRANSFER (BY
 CC FT SITE 267 267 INVOLVED IN AUTO-CLEAVAGE (BY
 CC FT ACT_SITE 270 270 ESSENTIAL FOR AUTO-CLEAVAGE (BY
 CC FT BINDING 197 197 CHOLESTEROL (BY SIMILARITY).
 CC FT SEQUENCE 418 AA; 46402 MW; CF000AFFFDF2F5795 CRC64;
 CC
 CC Query Match 58.6%; Score 1274; DB 1; Length 418;
 CC Best Local Similarity 61.2%; Pred. No. 1.3e-99;
 CC Matches 254; Conservative 46; Mismatches 101; Indels 14; Gaps 4;
 CC
 CC QY 5 RLRLPLHLCVLLVLLVPAAMGCGPGRVSGRRPRKLVPLAYKOPSPVNPRTKIGAS 64
 CC DB 2 RLRLVLLVSLTLSTLVV-SGLACGPGGCG-RRRHKKLTPLAYKOPIPVAVKRTIGAS 59
 CC QY 65 GRYGKTIARSSERKELTPNPNPDIIRKDEMTGADSLMORCDRLNSLAISYMNQWPG 124
 CC DB 60 GRYEKTIRNSERKELTPNPNPDIIRKDEMTGADSLMORCDRLNSLAISYMNQWPG 119
 CC QY 125 VKLRTGMDMDGHHSESLHYEGRAVDITTSDDRNKYGGLLRLAYAGFDWYYESKA 184
 CC DB 120 VKLRTGMDMDGHHSESLHYEGRAVDITTSDDRNKYGGLLRLAYAGFDWYYESKA 179
 CC QY 185 HVHGSVKSSESAAKKTGGCPAGQVRLSGARVALSAVRGDRVLAMGEDGSPFESDVL 244
 CC DB 180 HHSVKAKENSVAAKSGGCGPAGQVRLSGARVALSAVRGDRVLAMGEDGSPFESDVL 239
 CC QY 245 IFLREPRRLAPQVITPDPRRLATLPAHLFTADNHPARFATPSHQPOGOY 304
 CC DB 240 MFTDSDSTRVVFVIEQPEVEKITTLAHLFLVDNSTDHTMTAAASVSRQOKV 299
 CC QY 305 LVAGVPG-LQPARVAATHVAGAPLTKGTLVVEDVVAASCFAAVADHHLQLFWP 363
 CC DB 300 MVVDSGLKSKVIVORITTEGRGSPFAPVYRAGTIIVDRILASCIYAVIEDOGLAHAFAP 359
 CC QY 364 LRLFH-----SLAWGWTPEGVHWYPOLLYRLGLHLLIEGSGFPLGNS 407
 CC DB 360 ARLYYVSFLFPPONSSSRNATLQGGVHWYSRLYQMGWGLDLSNMLHPLGNS 414
 CC
 CC RESULT 7
 CC SHH_CHICK STANDARD; PRT; 425 AA.
 CC AC Q91035;
 CC DT 15-JUL-1999 (Rel. 38, Created)
 CC DT 15-JUL-1999 (Rel. 38, Last sequence update)
 CC DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE SONIC HEDGEHOG PROTEIN PRECURSOR (SHH).
GN SHH.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
OC Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LIMB BUD.
RX MEDLINE: 94094333.
RA Riddle R.D., Johnson L., Laufer E., Tabin C.;
RT "Sonic hedgehog mediates the polarizing activity of the ZPA."
RL Cell 75:1401-1416(1993).
RN [2]
RP FUNCTION AND AUTOPROTEOLYTIC CLEAVAGE.
RX MEDLINE: 95254654.
RA Roelink H., Porter J.A., Chiang C., Tanabe Y., Chang D.T.,
RA Beachy P.A., Jessell T.M.;
RT "Floor plate and motor neuron induction by different concentrations of
RT the amino-terminal cleavage product of sonic hedgehog
RT autoproteolysis."
RL Cell 81:445-455(1995).
CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN
CC ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION
CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE
CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER
CC TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A
CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED
CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE
CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE
CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH
CC FLOOR PLATE- AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD
CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS
CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION.
CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
CC TO MEDIANE SIGNALING TO OTHER CELLS (BT SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE POSTERIOR LIMB BUD
CC MESENCHYME, THE HENSEN'S NODE, THE NOTOCHORD, AND THE FLOOR PLATE
CC OF THE NEURAL TUBE.
CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE AT STAGE 17 DURING THE
CC INITIATION OF LIMB BUD FORMATION. FROM THAT POINT ONWARDS, THE
CC EXPRESSION PATTERN EXACTLY MATCHES THE LOCATION OF THE ZONE OF
CC POLARIZING ACTIVITY (ZPA).
CC -1- INDUCTION: BY RETINOID ACID.
CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
CC C-PRODUCT HAS NO SIGNALING ACTIVITY.
CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L28099; AAA72428.1; -
DR PRAM; PF01079; Hint; 1.
DR PRAM; PF01085; HH_Signal; 1.
DR PRINTS; PR00632; SONICHHOG.
KM Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
KW Signal.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 425 SONIC HEDGEHOG PROTEIN.
FT CHAIN 27 200 SONIC HEDGEHOG PROTEIN N-PRODUCT.

FT CHAIN 201 425
FT SITE 200 201
FT SITE 200 201
FT SITE 246 246
FT SITE 246 246
FT SITE 270 270
FT ACT_SITE 273 273
FT BINDING 200 200
FT DOMAIN 390 393
SQ SEQUENCE 425 AA; 46474 MW; DA9627443DA40173 CRC64;
Query Match 58.4%; Score 1270.5; DB 1; Length 425;
Best Local Similarity 63.0%; Pred. No. 2,6e-99;
Matches 257; Conservative 38; Mismatches 96; Indels 17; Gaps 5;
QY 20 LVPAANGCGPGRVYSGRRPRRLVPLAKQSPNVPKTLGASRGYRGIANSSEK 79
DB 19 LVSSGLTCGPGRGIG-KRRHPKRLPLAKQFIPNVAEKTLGASRGYRGIANSSEK 77
QY 80 ELTPNYPDITFKDEMTGADRLTORCKDRLNSATSVNMQGVKLRVTEGDEGHH 139
DB 78 ELTPNYPDITFKDEMTGADRLTORCKDRLNSATSVNMQGVKLRVTEGDEGHH 137
QY 140 SEESLHTEGRAVDITTSDRNRKYGTLRLAVEAFDPMYVESKARHCVSEKSAK 199
DB 138 SEESLHTEGRAVDITTSDRNRKYGTLRLAVEAFDPMYVESKARHCVSEKSAK 197
QY 200 TGGCPGAGVRLSAGARVALSAVRPDGVLMGDESPFTSDVLTLDREPHRLAPQV 259
DB 198 SGGCFPSATVHEHGCTKLVKDLSPGDRVLAADADRLIYSDLTLDREDSRKLFYV 257
QY 260 IETDPPRLALFPAHLFTADNTEPA-----REFATPSHVPQGVYVAVGPGQ-- 314
DB 258 IETDPPRLALFPAHLFTADNTEPA-----REFATPSHVPQGVYVAVGPGQ-- 317
QY 314 PARVAAYS-THVALGAYAPLTKHGTLVVEDVASCFAVADHHLAQLAFPLRFLSHLAW 372
DB 318 PASVHSVLSREASGAYAPLTAQGTILLNRLLACVAVIEHSHANAFAPRLAQGLLA 377
QY 373 G-----SWTPEGVHTVPLLRLGLLEGGFPHLGSAGS 411
DB 378 ALCPDGAIPPLATTTTGIHWYSRLYRIGSVWLDGDLHPLGMVAPAS 425
RESULT 8
SHH_HUMAN STANDARD; PRT; 462 AA.
AC Q15465;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE SONIC HEDGEHOG PROTEIN PRECURSOR (SHH) (HNG-1).
GN SHH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-FETAL LUNG;
RX MEDLINE: 96070431.
RA Marigo V., Roberts D.J., Lee S.M.K., Tsukurov O., Levi T.,
RA Gaster J.M., Epstein D.J., Gilbert D.J., Copeland N.G., Seidman C.E.,
RA Jenkins N.A., Seidman J.G., McMahon A.P., Tabin C.;
RT "Cloning, expression, and chromosomal location of SHH and IHH: two
RT human homologues of the Drosophila segment polarity gene hedgehog."
RL Genomics 28:44-51(1995).
RN [2]
RP SEQUENCE OF 1-187 FROM N.A.
RA Strong C., Graves T., Sutterer C., Ozerzky P.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [3]

RA Takahashi T., Takahashi T.C., Inoue K., Ogawa M., Takeshima K.:
 RT "Activation of two Cynops genes, fork head and sonic hedgehog, in
 RL animal cap explants.";
 CC Blochem. Biophys. Res. Commun. 218:395-401(1996).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT AND METAMORPHOSIS. INVOLVED
 CC IN LIMB FORMATION, PATTERNING OF THE CENTRAL NERVOUS SYSTEM AND
 CC VENTRAL SOMITE DIFFERENTIATION. INDUCES ECTOPIC CEMENT GLAND
 CC FORMATION IN EMBRYOS. BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH
 CC FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE
 CC TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC
 CC REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- INDUCTION: ACTIVATED BY ACTIVIN, BASIC FIBROBLAST GROWTH FACTOR
 CC (bFGF) AND FGF.
 CC -1- PFM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; D63339; BAA09657.1; -;
 DR PFM; PFO1079; HInt; 1.
 DR PFM; PFO1085; HR.Signal; 1.
 DR PRINTS; PR00632; SONICHOG.
 KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 KW Signal.
 FT SIGNAL. 1 26
 FT CHAIN 27 432
 FT CHAIN 27 200
 FT CHAIN 201 432
 FT SITE 200 201
 FT SITE 268 268
 FT SITE 268 268
 FT ACT_SITE 271 271
 FT BINDING 200 200
 FT SEQUENCE 432 AA; 47847 MW; B453C7E746C8E5A8 CRC64;
 SO
 Query Match 58.1%; score 1263; DB 1; length 432;
 Best Local Similarity 60.4%; Pred. NO. 1-le-98;
 Matches 252; Conservative 48; Mismatches 89; Indels 28; Gaps 6;
 QY 20 LVPAAWCGGGRVVGSRPRPKLVLAYVQSPNVEKTLGASRGYKARSSEK 79
 DB 19 LIVESGLSCGGRGIGTKRR-FKRLTPLAYKQFTPNVEKTLGASRGYKARSSEK 77
 QY 80 ELTPYNDIIFKDEENGGADRLMTQCKDRNSLAISVMNQWPKVRLVTEGMEDGH 139
 DB 78 ELTPYNDIIFKDEENGGADRLMTQCKDRNSLAISVMNQWPKVRLVTEGMEDGH 137
 QY 140 SEESLHYGRAVDITTSQDRNRKYLGLARLAEAGFDWYYSKAAHVCSSESAAK 199
 DB 138 FEESLHYGRAVDITTSQDRNRKYLGLARLAEAGFDWYYSKAAHVCSSESAAK 197

QY 200 TGCGPAGVAGVRLSESGARVALASVAPGDRVLAMGEDSPTEFSDVLTFLDEPHRLRAFOV 259
 DB 198 SGCGPAGVAGVRLSESGARVALASVAPGDRVLAMGEDSPTEFSDVLTFLDEPHRLRAFOV 257
 QY 260 IETODPPRLRLTAHLFTFADN-----TEPARFRTAFSHVPGQYVLVAG 308
 DB 258 IETS--KERVLTAAHLFLVQALPAGDSGDFSVSGSAGFRMFASVAGRVLTVD 315
 QY 309 VP--GLQPARVAASVTHVALGAVAPLTKHGLVVEDVAVSCFAAVADHHLAQLAEFWPLRL 366
 DB 316 REGGLREATVERVLEATGAVAPVTAHGVIVDRVLASCYAVIEHSHMAFAFLRY 375
 QY 367 -----FHSIANGSWP-----GEVHYRQQLYRGRLLLEGSPFLMGSGAGS 411
 DB 376 GLGILSFSPDQYSSHPAPQSGEVMWSEILYRIGTAVLODTHPIGMARKSS 432
 RESULT 10
 SHH_XENLA STANDARD; PRT; 444 AA.
 ID SHH_XENLA
 AC Q92000; Q91894;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-FEB-2000 (Rel. 39, Last annotation update)
 DE SONIC HEDGEHOG PROTEIN PRECURSOR (X-SHH) (VHH-1).
 GN SHH.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
 OC Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae; Xenopodinae;
 OC Xenopus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-INTESINE;
 RX MEDLINE; 95357169.
 RA Stolor M.A., Shi Y.-B.;
 RT "Xenopus sonic hedgehog as a potential morphogen during embryogenesis
 RT and thyroid hormone-dependent metamorphosis.";
 RL Nucleic Acids Res. 23:2555-2562(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-EMBRYO;
 RX MEDLINE; 95401852.
 RA Ekker S.C., McGrew L.L., Lai C.-J., Lee J.-J., von Kessler D.P.,
 RA Moon R.T., Beachy P.A.;
 RT "Distinct expression and shared activities of members of the hedgehog
 RT gene family of Xenopus laevis.";
 RL Development 121:2337-2347(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-NOTOCHORD;
 RX MEDLINE; 96028338.
 RA Ruiz I Altaba A., Jessell T.M., Roelink H.;
 RT "Restrictions to floor plate induction by hedgehog and winged-helix
 RT genes in the neural tube of frog embryos.";
 RL Mol. Cell. Neurosci. 6:106-121(1995).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT AND METAMORPHOSIS. INVOLVED
 CC IN LIMB FORMATION, PATTERNING OF THE CENTRAL NERVOUS SYSTEM AND
 CC VENTRAL SOMITE DIFFERENTIATION. INDUCES ECTOPIC CEMENT GLAND
 CC FORMATION IN EMBRYOS. BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH
 CC FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE
 CC TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC
 CC REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN NOTOCHORD AND NEURAL
 CC FLOOR PLATE DURING EMBRYOGENESIS. IN TADPOLE, HIGH EXPRESSION IS
 CC OBSERVED IN PANCREAS/STOMACH, MODERATE EXPRESSION IN TAIL, AND LOW
 CC EXPRESSION IN INTESTINE, BRAIN, AND HIND LIMB.
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED AT THE NEURULA (STAGES 16-17).

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: L27340; AAA20999.1; -
DR PFM: PF01079; Hint: 1.
DR PFM: PF01085; HH_Signal; 1.
DR PRINTS: PR00632; SONICHOG.
KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
KW Signal.

FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 437 SONIC HEDGEHOG PROTEIN.
FT CHAIN 25 198 SONIC HEDGEHOG PROTEIN N-PRODUCT.
FT CHAIN 199 437 SONIC HEDGEHOG PROTEIN C-PRODUCT.
FT SITE 198 199 CLEAVAGE (AUTO-).
FT SITE 244 244 INVOLVED IN CHOLESTEROL TRANSFER (BY
FT SITE 268 268 INVOLVED IN AUTO-CLEAVAGE (BY
FT ACT_SITE 271 271 ESSENTIAL FOR AUTO-CLEAVAGE (BY
FT BINDING 198 198 SIMILARITY).
FT DOMAIN 383 383 CHOLESTEROL (BY SIMILARITY).
FT SEQUENCE 437 AA; 47630 MW; 0DBFC19F0D1662A0 CRC64;

Query Match 57.3%; Score 1246.5; DB 1; Length 437;
Best Local Similarity 59.9%; Pred. No. 2.7e-97;
Matches 257; Conservative 40; Mismatches 99; Indels 37; Gaps 8;

QY 13 CLVLL---LTVPAWCGPGRVVGSRPRKLVPLAYKQFSPNVEKTLGASGRREG 69
DB 8 CELVALASSLLVCPGL-ACGPGRGF-KRQHPKLTPLAYKQFIPNVAEKTLAGSGREG 65
QY 70 KIARSEFEKELTPNPNNDIIFKDEENTGADRLMTORCKDRNLASLISVMNMPGVKRLV 129
DB 66 KITRSEFEKELTPNPNNDIIFKDEENTGADRLMTORCKDRNLASLISVMNMPGVKRLV 125
QY 130 TEGWDEGHSESLHYEGRAVDITSDRNRKYGILLALAEAGPDWYYSKAVHCS 189
DB 126 TEGWDEGHSESLHYEGRAVDITSDRNRKYGILLALAEAGPDWYYSKAVHCS 185
QY 190 VKSEHSAAKTGCGEPAGAVRLSAGARVALSARPGDRVLAMGDSPTFSDVLIIFLDR 249
DB 186 VKAENSVAKSDGCPGSAKTVHLEOGGKTVKDLSPGDRVLADDOGRLLYSDFLFLDR 245
QY 250 EPHRLAROVITODPPRLALTPAHLFTA-DNHTEPARRATPASHVQGVYLVAG 308
DB 246 DEGAARKVYVITREPREKRLTLTAHLLEFVAPHNSGPTPGSPPLFASRVRGORYVVA 305
QY 309 VGG---LQPARVAVS-THVALGAVAPLTGHTLVDEVVASCFAVAVDHHLAOLAFWP 363
DB 306 EGGGRRLPLPAVHSTVLAEEAAGATPILADGTIILNNVLSCAVIVIEHSHAAAFAP 365
QY 364 LRLHSL-----AMGSWPEGEVHWYPLLRYRLGLLEE 398
DB 366 FLALHALALAPARTDGGGGSIPAPGSVAEARGAPPA-GIHWYSOLLYHIGTWLDS 424
QY 399 GSFHPLGMS 407
DB 425 ETLHPLGMA 433

RESULT 12
TWHH_BRARE STANDARD; PRT: 416 AA.
AC Q90419;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE TIGGY-WINKLE HEDGEHOG PROTEIN PRECURSOR (TWHH).
GN TWHH.

OS Brachydanio rerio (Zebrafish) (Zebra danio).
CC Euarystota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
CC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
CC Cyprinoidae; Cyprinidae; Rasbora; Danio.
RN [1]

RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
RC TISSUE-EMBRYO;
RX MEDLINE; 96014264.
RA Ekker S.C., Ungar A.R., Greenstein P., von Kessler D.P.,
RA Porter J.A., Moon R.T., Beachy P.A.;
RT "Patterning activities of vertebrate hedgehog proteins in the
RT developing eye and brain";
RL Curr. Biol. 5:944-955(1995).

-1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
PATTERNING EVENTS DURING DEVELOPMENT. INVOLVED IN DORSO-VENTRAL
PATTERNING OF THE BRAIN AND IN EARLY PATTERNING OF THE DEVELOPING
EYES.

-1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).

-1- TISSUE SPECIFICITY: EXPRESSED IN THE VENTRAL MIDLINE OF THE NEURAL
TUBE AND BRAIN. IN THE DEVELOPING BRAIN, EXPRESSION OCCURS IN
DOMAINS THAT INCLUDE A DISCRETE REGION IN THE FLOOR OF THE
DIENCEPHALON. NOT DETECTED IN THE NOTOCHORD OR DEVELOPING FIN BUD.

-1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY
AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).

-1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U30710; AAC59741.1; -
DR ZFIN: ZDB-GENE-980526-41; TWHH.
DR PFM: PF01079; Hint: 1.
DR PRINTS: PR00632; SONICHOG.
DR PFM: PF01085; HH_Signal; 1.
KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
KW Signal.

FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 416 TIGGY-WINKLE HEDGEHOG PROTEIN.
FT CHAIN 27 200 TIGGY-WINKLE HEDGEHOG PROTEIN N-PRODUCT.
FT CHAIN 201 416 TIGGY-WINKLE HEDGEHOG PROTEIN C-PRODUCT.
FT SITE 200 201 CLEAVAGE (AUTO-).
FT SITE 270 270 INVOLVED IN AUTO-CLEAVAGE (BY
FT ACT_SITE 273 273 ESSENTIAL FOR AUTO-CLEAVAGE (BY
FT BINDING 200 200 CHOLESTEROL (BY SIMILARITY).
FT SEQUENCE 416 AA; 46576 MW; 61EC2218309CFE59 CRC64;

Query Match 56.4%; Score 1226.5; DB 1; Length 416;
Best Local Similarity 58.2%; Pred. No. 1.2e-95;
Matches 241; Conservative 53; Mismatches 99; Indels 21; Gaps 6;

QY 9 RLHF---CLVLLLLVPAWCGPGRVVGSRPRKLVPLAYKQFSPNVEKTLGAS 64
DB 4 RLHMFPLHCLISLLPLPCGACGPRGCG-KRHHPKLTPLAYKQFIPNVAEKTLAGS 62
QY 65 GRKGIARSSERKELTPNPNNDIIFKDEENTGADRLMTORCKDRNLASLISVMNMPG 124

DB 63 GKTEGKTRNSEKELTPYNDIIFKDEENTGADRLMTRKDKLNLSIAISVMNMPG 122
125 VKLRVTEGMDDEGHSESLHYEGRAVDITTSDRNKKYGLARLAVEAGFDWVYESKA 184
123 VKLRVTEGMDDEGHSESLHYEGRAVDITTSDRNKKYGLARLAVEAGFDWVYESKA 182
185 HVCYSKSESAAKTGCGCPAGAOVRLESGARVALSAPGGRVLAAMEGDSPTSDVL 244
183 HHCYSKSESAAKTGCGCPAGAOVRLESGARVALSAPGGRVLAAMEGDSPTSDVL 242
245 IFDREPHRLRAQVLETODPPRRLALTPAHLEFADNHTEPARFATPASHVOGQVY 304
243 MFDHDPTRKQVLETSEPTFKLITAAHLVY--GNSAASGATATPASNKVEDYV 300
DB 301 LVWEDCESLKSATYKRIYEEHSGSPAYTAGITIVDQVLAACVAVIENHKAHMAFA 360
QY 305 LV--AGVPGLOPARVAVSTHVALGAYAPLTKGTLVEDVAVASCPAAVDHHLAQLAFW 362
DB 301 LVWEDCESLKSATYKRIYEEHSGSPAYTAGITIVDQVLAACVAVIENHKAHMAFA 360
QY 363 PLALFSLANGSWT-----PGEVHWYPOLYRLGRLLLESGSFHPLGM 406
DB 361 PVRLCRKLK--TWLFPARESNNFQEDGIMHYSNMLFHIGSWLLDRDSFHLGI 412
RESULT 13
ID IHH_BRARE STANDARD: PRT: 412 AA.
AC Q98862;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE INDIAN HEDGEHOG PROTEIN PRECURSOR (IHH) (ECHIDNA HEDGEHOG PROTEIN).
GN IHH OR EHH.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinidae; Rasbora; Danio.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 96310864.
RA Currie P.D., Ingham P.W.;
RT "Induction of a specific muscle cell type by a hedgehog-like protein
in zebrafish."
RL Nature 382:452-455(1996).
RN [2]
RP SEQUENCE OF 113-170 FROM N.A.
RX TISSUE-MUSCLE;
RA Zardoya R., Abouheif E., Meyer A.;
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
closely related to the zebrafish."
RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE
NOTOCHORD THAT INDUCES SOMITE PATTERNING AND MUSCLE PIONEER
DIFFERENTIATION.
CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN THE NOTOCHORD.
CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE AT THE MID-GASTRULA STAGE.
CC -1- DISAPPEARS AT THE END OF THE SOMITOGENESIS.
CC -1- P.TM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
CC or send an email to license@sdb.ch).
CC
DR EMBL: Y08426; CAA69702.1; -.
DR ZFIN: ZDB-GENE-980526-135; EHH.
DR PFM: PF01079; Hint: 1.
DR PFM: PF01085; HH_Signal: 1.
DR PRINTS: PR00632; SONTCHOG.
KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
KW Signal.
FT SIGNAL 1 23
FT CHAIN 1 412
FT CHAIN 24 197
FT CHAIN 198 412
FT SITE 197 198
FT SITE 197 198
FT SITE 246 246
FT SITE 269 269
FT SITE 272 272
FT ACT_SITE 272 272
FT BINDING 197 197
FT CONFLICT 121 121
SQ SEQUENCE 412 AA: 45746 MW: 2104350528278CAL CRC64:
Query Match 55.0%; Score 1196; DB 1; Length 412;
Best Local Similarity 61.7%; Pred. No. 4.4e-93;
Matches 258; Conservative 35; Mismatches 99; Indels 32; Gaps 13;
QY 9 RHFCVLV---LLVVPAMGCGPGRVSGRRPRKRVPLAYKQSPNVYEKLTGASG 65
DB 2 RUSTAAALTGFLAFSPAYDGGPGRGYG-KRRPRKRLPLAYKQSPNVYEKLTGASG 60
QY 66 RYEGKTIARSSERKELTPYNDIIFKDEENTGADRLMTRKDKLNLSIAISVMNMPGV 125
DB 61 RYEGKTIARSSERKELTPYNDIIFKDEENTGADRLMTRKDKLNLSIAISVMNMPGV 120
QY 126 VKLRVTEGMDDEGHSESLHYEGRAVDITTSDRNKKYGLARLAVEAGFDWVYESKAH 185
DB 121 VKLRVTEGMDDEGHSESLHYEGRAVDITTSDRNKKYGLARLAVEAGFDWVYESKAH 180
QY 186 VHCYSKSESAAKTGCGCPAGAOVRLESGARVALSAPGGRVLAAMEGDSPTSDVL 244
DB 181 VHCYSKSESAAKTGCGCPAGAOVRLESGARVALSAPGGRVLAAMEGDSPTSDVL 240
QY 243 VLIFDREPHRLRAQVLETODPPRRLALTPAHLEFADNHTEPARFATPASHVOGQVY 304
DB 241 VLIFDREPHRLRAQVLETODPPRRLALTPAHLEFADNHTEPARFATPASHVOGQVY 300
QY 298 VDPGQVLYVAGPGLAPARVAVSTHVAL---GAYAPLTKGTLVEDVAVASCPAAVDH 353
DB 300 AVGQCCILL---GKLRKRFSDI--THVGREDQGLPPYTAGITIVDQVLAACVAVI 354
QY 354 HHQAQLAFPLFLFSLANGSWT-PGE---GVHWYPOLYRLGRLLLESGSFHPLGM 406
DB 355 QRLAHAFAPRLRLY---SWTGPOVLKNGLHWYSOVLIGIKLLDSELFLPLAL 407
RESULT 14
ID DHH_MOUSE STANDARD: PRT: 396 AA.
AC Q61488;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE DESERT HEDGEHOG PROTEIN PRECURSOR (DHH) (HHS-3).

GN DHH.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-129/SV.
 RX MEDLINE: 94094334.
 RA Echelard Y., Epstein D.J., St Jacques B., Shen L., Mohler J.,
 RA McMahon J.A., McMahon A.P.,
 RT "Sonic hedgehog, a member of a family of putative signaling
 RT molecules, is implicated in the regulation of CNS polarity.";
 RL Cell 75:1417-1430(1993).
 RN [2]
 RP SEQUENCE OF 120-168 FROM N.A.
 RX MEDLINE: 95236997.
 RA Zhang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,
 RA Zhao R., Seldin M.F., Fallon J.F., Beachy P.A.;
 RT "Products, genetic linkage and limb patterning activity of a murine
 RT hedgehog gene.";
 RL Development 120:3339-3353(1994).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT. MAY FUNCTION AS A
 CC SPERMATOCYTE SURVIVAL FACTOR IN THE TESTES. ESSENTIAL FOR TESTES
 CC DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ADULT TESTES. NOT EXPRESSED IN
 CC LIMB BUDS.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X76292; CAA53924.1; -
 DR MGD: MGI:94891; DHH.
 DR PFAM: PF01079; HInt: 1.
 DR PFAM: PF01085; HH: signal: 1.
 DR PRINTS: PRO0632; SONICHHOG.
 KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 KW signal.
 FT SIGNAL. 1 22
 FT CHAIN. 23 396
 FT CHAIN. 23 198
 FT CHAIN. 199 396
 FT SITE. 199 199
 FT SITE. 244 244
 FT SITE. 268 268
 FT ACT_SITE. 271 271
 FT BINDING. 198 198
 FT SEQUENCE. 396 AA; 43542 MW; AFE6051BE950FD8 CRC64;
 Query Match 54.0%; Score 1175; DB 1; Length 396;

Best Local Similarity 58.5%; Pred. No. 2.4e-91;
 Matches 234; Conservative 50; Mismatches 102; Indels 14; Gaps 5;
 QY 3 PALRLRLHFCVLLVLLVPAWAGCGPR-VYGSRRPRPRPKVPLAYKQFSPNVPEKTL 61
 DB 4 PASLPLIC--CLALLA-----SAQSGPGRPGVGRRIYRKQVLLYQFQPSMNERLT 57
 QY 62 GASRGYKTIANSSEKELTPNYNDIIFKDEENTGADRLMTQRCCKRLNSLAISVNNQ 121
 DB 58 GASGPAEGHVTGSEKFRPLVYNYNDIIFKDEENSGADRLMTQRCCKRLNSLAISVNNQ 117
 QY 122 WPGVKRLVTEGMDGHHDEEELHYEGRAVDITTSRDNRKKGILLARLVEGFDWVYVE 181
 DB 118 WPGVKRLVTEGMDGHHDEEELHYEGRAVDITTSRDNRKKGILLARLVEGFDWVYVE 177
 QY 182 SKAHVCSYKSEHSAAKGCGEPAGQVRLSEGAVALASAVRPGDVLAMGEDSGPTSS 241
 DB 178 SRNHIVSYKADNSLAVRAGCGEPGATVRLASGEKGLRELHGRGWVLAADAAGRVPT 237
 QY 242 DVILFIDREPHRLARFQVLETQDPPRLALTPAHLLFTADNHTPEARFRATFASHVQPG 301
 DB 238 PVLLFDRLQRRASFVAVETERPRLKLLTPMHLVFAARGPAPAGDPAPVFAFRLRAG 297
 QY 302 QVLYVAGVGLQPARVAASTHVALGAVAPLTKGTLVYEDVVASCFEAVADHHLAQLAF 361
 DB 298 DSVLAPGAGDALPARVARVAREAVGFAPLTAHGTLLVNDVASCYAVLESHQWHRAF 357
 QY 362 WPLRLFSLAMGSMTPG-----EGVHWYPLLRYRLGLLL 396
 DB 358 APRLRLHAL--GALDPGAVPFGMHWYRLRYRLAEELM 395
 RESULT 15
 DHH HUMAN STANDARD: PRT: 396 AA.
 ID 043323; Q15794;
 AC 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE DESERT HEDGEHOG PROTEIN PRECURSOR (DHH) (HHC-3).
 GN DHH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tate G., Endo Y., Mitsuura T.;
 RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 85-178 FROM N.A.
 RC TISSUE-KIDNEY;
 RA Drummond I.A.;
 RT "Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases."
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT. MAY FUNCTION AS A
 CC SPERMATOCYTE SURVIVAL FACTOR IN THE TESTES. ESSENTIAL FOR TESTES
 CC DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AB010994; BAA24866.1;
DR EMBL; AB010581; BAA24866.1; JOINED.
DR EMBL; AB010993; BAA24866.1; JOINED.
DR EMBL; U59748; AAB03398.1;
DR PFAM; PF01079; Hntc; 1.
DR PFAM; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHHOG.
KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
KW Signal.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 396 DESERT HEDGEHOG PROTEIN.
FT CHAIN 23 198 DESERT HEDGEHOG PROTEIN N-PRODUCT.
FT CHAIN 199 396 DESERT HEDGEHOG PROTEIN C-PRODUCT.
FT SITE 198 199 CLEAVAGE (AUTO-) (BY SIMILARITY).
FT SITE 244 244 INVOLVED IN CHOLESTEROL TRANSFER (BY
FT SITE 268 268 SIMILARITY).
FT SITE 268 268 INVOLVED IN AUTO-CLEAVAGE (BY
FT ACT_SITE 271 271 SIMILARITY).
FT BINDING 198 198 ESSENTIAL FOR AUTO-CLEAVAGE (BY
FT CONFLICT 177 177 SIMILARITY).
FT CONFLICT 177 177 CHOLESTEROL (BY SIMILARITY).
SQ SEQUENCE 396 AA; 43577 MW; FCE4FB21972C3AD5 CRC64; E -> G (IN REF. 2).

Query Match 53.9%; Score 1171.5; DB 1; Length 396;
Best Local Similarity 59.9%; Pred. No. 4.8e-91;
Matches 233; Conservative 47; Mismatches 100; Indels 9; Gaps 4;

QY 14 LVLLLVVPAAMGCGRGR-VYGSRRRPRKLVPLAYKQSPNPEKTLGASGRYEGKIA 72
DB 10 LCCLALLPAAO-SCGPGRPGVRRRYAKQLVPLKYQFVGPVVERTLGASGPAEGRVA 68
QY 73 RSERFKELTPNYPNDITFKDENTGADRLTORCKDRNLSTISVMNQMPGVLRVTEG 132
DB 69 RGSERFRLVYNYPNDITFKDENSGADRLMTERCKERNALAIYANMMPGVRLRVTEG 128
QY 133 WDEGHHSESLHYEGRAVDITTSDRDNKYGLLARLAVEAGFDMVYYESKAHYCSYKS 192
DB 129 WDEGHHHODSLHYEGRALDITTSDRDNKYGLLARLAVEAGFDMVYYESRNHVSYKA 188
QY 193 EHSAAKTGCGCFPAAGVRLSAGARVALSAVRPGDRLVLAAGEDSPFSDVLIFLDRPH 252
DB 189 DNSTLAVRAGCGFPAGATVRLMSGERKGLRELHRGDVWLADASGRVPTPVLLFLDRDLO 248
QY 253 RLRAFOVETODPPRRLLTPAHLLETPADNHTPEAPRATFASHVOPGQVVLVAGVPG 312
DB 249 RRASTFVAETEMPRKLLTPMHLVFAARGPAPAGDPAPFARLRAGDSVLRAGDAL 308
QY 313 QPARVAAYSTHVALGAYAPLTKHGLVVEDVVASCEAAVADHHLAQLAFWPLRLFHSIAW 372
DB 309 RPARVAARVAREAVGVFAPLTAHGLLVNDVLAASYAVLSEHQWNAHRAFAPLRLHLAL-- 367
QY 373 GSWTPG-----EGYHWYQLLYRGLRL 396
DB 367 GALLPGAVOPTGMWYSRLYLAEELL 395

Search completed: June 5, 2000, 08:19:54
Job time: 158 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2000, 07:56:48 ; Search time 52.13 seconds
(without alignments)
2969.733 Million cell updates/sec

Title: US-08-900-220-8

Perfect score: 1191
Sequence: 1 ATGGCTCTCTGACCAATCT.....CGAGAGAGCTACTGGCTGA 1191

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 segs, 6492525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5C.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/5D.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/6.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq:*
7: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	967.6	81.2	US-08-176-427B-3	Sequence 3, Appl1
2	967.6	81.2	US-08-356-060A-2	Sequence 2, Appl1
3	419.8	35.2	US-08-176-427B-7	Sequence 7, Appl1
4	419.8	35.2	US-08-356-060A-4	Sequence 4, Appl1
5	407	34.2	PCT-US95-02315-1	Sequence 1, Appl1
6	390.6	32.8	US-08-356-060A-6	Sequence 6, Appl1
7	389	32.7	US-08-748-581-5	Sequence 5, Appl1
8	389	32.7	US-08-748-591-10	Sequence 10, Appl1
9	381.4	32.0	US-08-176-427B-1	Sequence 1, Appl1
10	381.4	32.0	US-08-356-060A-5	Sequence 5, Appl1
11	355.8	29.9	US-08-176-427B-5	Sequence 1, Appl1
12	355.8	29.9	US-08-356-060A-3	Sequence 3, Appl1
13	355.8	29.9	US-08-356-060A-7	Sequence 7, Appl1
14	355.8	29.9	US-08-176-427B-9	Sequence 9, Appl1
15	316.2	26.5	US-08-356-060A-5	Sequence 5, Appl1
16	316.2	26.5	PCT-US95-15463-2	Sequence 2, Appl1
17	101.4	8.5	PCT-US95-15923-2	Sequence 2, Appl1
18	99.8	8.4	PCT-US95-15463-1	Sequence 1, Appl1
19	99.8	8.4	PCT-US95-15923-1	Sequence 1, Appl1
20	44.6	16.42	US-08-232-015-2	Sequence 2, Appl1
21	44.6	16.42	US-08-648-657-14	Sequence 14, Appl1
22	43	3.6	US-08-073-384C-3	Sequence 3, Appl1
23	43	3.6	US-08-254-359A-3	Sequence 3, Appl1
24	43	3.6	US-08-483-043-3	Sequence 3, Appl1
25	43	3.6	US-08-481-238-3	Sequence 3, Appl1
26	43	3.6	US-08-471-066B-3	Sequence 3, Appl1
27	43	3.6	US-08-484-956-3	Sequence 3, Appl1

C 28	43	3.6	2504	3	US-08-757-653-3	Sequence 3, Appl1
C 29	43	3.6	2504	3	US-08-599-491-3	Sequence 3, Appl1
C 30	43	3.6	2504	4	US-08-756-386-3	Sequence 3, Appl1
C 31	43	3.6	2504	4	US-08-823-516-3	Sequence 3, Appl1
C 32	43	3.6	2504	5	US-08-682-853A-3	Sequence 3, Appl1
C 33	43	3.6	2505	1	US-07-977-434-9	Sequence 9, Appl1
C 34	43	3.6	2505	2	US-08-458-819-9	Sequence 9, Appl1
C 35	43	3.6	2505	6	PCT-US91-07035-9	Sequence 9, Appl1
C 36	43	3.6	2640	1	US-08-384-490-30	Sequence 30, Appl1
C 37	43	3.6	2640	2	US-08-459-383-30	Sequence 30, Appl1
C 38	42.4	3.6	4257	3	US-08-680-473-1	Sequence 1, Appl1
C 39	42.4	3.6	9757	1	US-08-093-453B-1	Sequence 1, Appl1
C 40	42.4	3.6	9759	1	US-08-459-041A-1	Sequence 1, Appl1
C 41	42.4	3.6	12001	2	US-08-458-568A-11	Sequence 11, Appl1
C 42	41.6	3.5	20235	2	US-07-642-734C-3	Sequence 3, Appl1
C 43	41.6	3.5	30001	1	US-08-125-468-1	Sequence 1, Appl1
C 44	41.6	3.5	30001	3	US-08-474-933-1	Sequence 1, Appl1
C 45	41.4	3.5	1794	6	PCT-US95-14418-3	Sequence 3, Appl1

ALIGNMENTS

```
RESULT 1
US-08-176-427B-3
: Sequence 3, Application US/08176427B
: Patent No. 5789543
:
: GENERAL INFORMATION:
: APPLICANT: Ingham, Phillip W.
: APPLICANT: McMahon, Andrew P.
: APPLICANT: Tablin, Clifford J.
: TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
: TITLE OF INVENTION: Proteins and Uses Related Thereto
: NUMBER OF SEQUENCES: 33
: CORRESPONDENCE ADDRESS:
: ADDRESS: LAHIVE & COCKFIELD
: STREET: 60 State Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: ASCII(text)
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/176,427B
: FILING DATE: 30-DEC-1993
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Vincent, Matthew P.
: REGISTRATION NUMBER: 36,709
: REFERENCE/DOCKET NUMBER: HMT-006
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 227-7400
: TELEFAX: (617) 227-5941
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1190 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1191
:
: US-08-176-427B-3
:
: Query Match 81.2%; Score 967.6; DB 2; Length 1190;
: Best Local Similarity 88.3%; Pred. No. 5e-194;
: Matches 1051; Conservative 0; Mismatches 139; Indels 0;
```



```

QY 401 ACCAGCTCAGGATTCACCTCCACCTACGAAAGCCGCTTGTGACATCACTACGTGTGAC 460
DB 401 ATCATTCAGAGAGAGTCTTACACTATGAGGGTCGAGAGATGACATCACACAGCTCCGACC 460
QY 461 GCGACCGCAACAAGATGAGGTGTGCGGCGCGCTCGACAGTGAAGCCGGCTCGACTGGG 520
DB 461 GCGACCGCAACAAGATGAGGTGTGCGGCGCGCTCGACAGTGAAGCCGGCTCGACTGGG 520
QY 521 TCTACTAGAGATCCCGCAACACAGCTCCAGTGTGCTGCTCAAGCTATACACTACTGCGG 580
DB 521 TCTACTAGAGATCCCGCAACAGCTCCAGTGTGCTGCTCAAGCTATACACTACTGCGG 580
QY 581 TCCGGGGCGGGGCTGCTTCCGGGAAATGCAACTGTGGCTGTGAGACGGCGGACGGA 640
DB 581 CCAAAATCCGGGCTGCTTCCGGGAAATGCAACTGTGGCTGTGAGACGGCGGACGGA 640
QY 641 AAGGCTGCGGAGACTGCAACCGGAGACTGTGTTGGCGCCATGCGTAGGCGCGG 700
DB 641 AGCTGTGAGAGACTTACGTCCCGGAGACCGGCTGTGCGGCTGACGACGAGCGCGG 700
QY 701 TGTGCCCCAGCCGCTGCTGCTCTTCTGTGACCGGAGACTTGTGAGCGCGGCTTCAATTG 760
DB 701 TGTGCCCCAGCCGCTGCTGCTCTTCTGTGACCGGAGACTTGTGAGCGCGGCTTCAATTG 760
QY 761 TGGCTGTGAGAGACGAGTGGCTCCACGCAACTGTGCTACGCGCTGTGACCTGTGT 820
DB 761 ACGTGTGAGAGACGAGTGGCTCCACGCAACTGTGCTACGCGCTGTGACCTGTGT 820
QY 821 TTGCGGCTC-----GAGGCGCGCGCGCGCGCGCGCGAGCTTGTGACCGGTGT 868
DB 821 TGTGCGCGCGCGCGCAACAGACTGTGGGGCGCGCGCGCGCGCGCGCGCTTGTGACCG 880
QY 869 TCGCGCGCGCGCTACGCGCTGTGGGAGCTCGGTGTGCGCGCGCGCGCGGGA--TGGCTTC 925
DB 869 TCGCGCGCGCGCTACGCGCTGTGGGAGCTCGGTGTGCGCGCGCGCGCGGGA--TGGCTTC 925
QY 881 GCGTGTGCGCGCGCGCGCGCGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 940
DB 881 GCGTGTGCGCGCGCGCGCGCGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 940
QY 926 GCGCGCGCGCGCTGCGCGCTGTGCG--GCGGAGAGAGCGCTGTGCGCGCGCGCG 982
DB 926 GCGCGCGCGCGCTGCGCGCTGTGCG--GCGGAGAGAGCGCTGTGCGCGCGCGCG 982
QY 941 TCGCGCGCGCGCTGCAACGCTGTGAGCTGTGAGAGAGAGAGCGCGCGCGCTGCG 1000
DB 941 TCGCGCGCGCGCTGCAACGCTGTGAGCTGTGAGAGAGAGAGCGCGCGCGCTGCG 1000
QY 983 TCAACGCGCGCGAGAGAGCTGTGCTGTGAGAGAGAGCTGTGCTGTGCTGTGCTGTG 1042
DB 983 TCAACGCGCGCGAGAGAGCTGTGCTGTGAGAGAGAGCTGTGCTGTGCTGTGCTGTG 1042
QY 1001 TCAGCGCGCGAGAGAGCTGTGCTGTGAGAGAGAGCTGTGCTGTGCTGTGCTGTG 1060
DB 1001 TCAGCGCGCGAGAGAGCTGTGCTGTGAGAGAGAGCTGTGCTGTGCTGTGCTGTG 1060
QY 1043 AGAGTCAACAGTGTGCGAGAGCGCGCTTGTGCGCGCTGTGAGAGAGCGCGCTGAGG 1102
DB 1043 AGAGTCAACAGTGTGCGAGAGCGCGCTTGTGCGCGCTGTGAGAGAGCGCGCTGAGG 1102
QY 1061 AGAGTCAACAGTGTGCGAGAGCGCGCTTGTGCGCGCTGTGAGAGAGCGCGCTGAGG 1120
DB 1061 AGAGTCAACAGTGTGCGAGAGCGCGCTTGTGCGCGCTGTGAGAGAGCGCGCTGAGG 1120
QY 1103 CGGCTGT 1109
DB 1103 CGGCTGT 1127

```

RESULT 4

US-08-356-060A-4

Sequence 4, Application US/08356060A

Patent No. 5844079

GENERAL INFORMATION:

APPLICANT: Ingham, Phillip W.

APPLICANT: McMahon, Andrew P.

APPLICANT: Tablin, Clifford J.

TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing

NUMBER OF SEQUENCES: 47

CORRESPONDENCE ADDRESS:

ADDRESS: LAHIVE & COCKFIELD

STREET: 60 State Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,060A
FILING DATE: 14-DEC-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMT-006CP
TELEPHONE/COMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ. ID NO. 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1313 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1314
US-08-356-060A-4

Query Match 35.28; Score 419.8; DB 3; Length 1313;
Best Local Similarity 64.58; Pred. No. 1e-79;
Matches 701; Conservative 0; Mismatches 362; Indels 24; Gaps 4;

QY 41 CACTTCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 100
DB 47 GCGTGTGCTGTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 106
QY 101 GCTATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 160
DB 107 AC-----CCCAAAAGAGTGAACCGCTTTAGCGTACAGAGAGAGAGAGAGAGAGAG 160
QY 161 AGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 220
DB 161 AGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 220
QY 221 TCCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 280
DB 221 TTAAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 280
QY 281 GAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 340
DB 281 GAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 340
QY 341 TGAAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 400
DB 341 TGAAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 400
QY 401 ACCAGCTCAGGATTCACCTCCACCTACGAAAGCGCGCTGTGAGACATCACTACGCTGAC 460
DB 401 ATCATTCAGAGAGAGTCTTACACTATGAGGGTCGAGAGATGACATCACACAGCTCCGACC 460
QY 461 GCGACCGCAACAAGATGAGGTGTGCGGCGCGCTCGACAGTGAAGCCGGCTCGACTGGG 520
DB 461 GCGACCGCAACAAGATGAGGTGTGCGGCGCGCTCGACAGTGAAGCCGGCTCGACTGGG 520
QY 521 TCTACTAGAGATCCCGCAACAGCTCCAGTGTGCTGCTCAAGCTATACACTACTGCGG 580
DB 521 TCTACTAGAGATCCCGCAACAGCTCCAGTGTGCTGCTCAAGCTATACACTACTGCGG 580
QY 581 TCCGGGGCGGGGCTGCTTCCGGGAAATGCAACTGTGGCTGTGAGACGGCGGACGGA 640
DB 581 CCAAAATCCGGGCTGCTTCCGGGAAATGCAACTGTGGCTGTGAGACGGCGGACGGA 640

```

```

QY 641 AAGGCTGCGGGAATCAGCCGCGAGACTGGTTTGGCGCGCATGCTCAGCGCGG 700
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 641 AGCTGTGAAGACTTACGTCCCGAGACCGCGTGTGCGGCTAGACAGACCGCGCG 700
QY 701 TGTGTCGCGGCGGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 760
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 701 TGTGTACAGGACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 760
QY 761 TGTGTGTGAGACGAGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 820
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 761 ACGTGATCGAGACGCTGAGCGCGCGCGAGCGCTGTCTCTCTCTCTCTCTCTCT 820
QY 821 TTGCGCGCTC-----GAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGT 868
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 821 TCGTGGCGCGCGCACAGACTCGGCGCGCGCGCGCGCGCGCGCGCGCGCTGT 880
QY 869 TCGGCGCGCGCTGAGCGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 925
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 881 GCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGC 940
QY 926 GAGCGAGCGCGCTGCGCGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 982
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 941 TCGCGCGCGCGCTGAGCGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 1000
QY 983 TCACCGCGCGAGCGCGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT 1042
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1001 TCACGCGCGCGAGCGCGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 1060
QY 1043 AGAGTACACAGTGGCGCGCGCGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT 1102
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1061 AGAGCGCGAGCTGTGAGCGCGCGCGCTGTGAGCTGTGAGCTGTGAGCTGTG 1120
QY 1103 CGCTGCT 1109
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1121 CCGCGCT 1127

```

RESULT 5

PCT-US95-02315-1
Sequence 1, Application PC/TUS9502315

GENERAL INFORMATION:

APPLICANT: Jessell, Thomas M.

APPLICANT: Dodd, Jane

APPLICANT: Roelink, Henk

APPLICANT: Edlund, Thomas

TITLE OF INVENTION: DNA ENCODING A VERTEBRATE HOMOLOGY OF

TITLE OF INVENTION: HEDGEHOG, VHH-L, EXPRESSED BY THE NOTOCHORD, AND USES

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/02315

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: John P. White

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 45375-A-PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 88-0400

TELEFAX: (212) 391-0525

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1715 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 315..1628
PCT-US95-02315-1

Query Match 34.2%; Score 407; DB 6; Length 1715;
Best Local Similarity 63.8%; Pred. No. 4.8e-77;
Matches 693; Conservative 0; Mismatches 370; Indels 24; Gaps 4;

```

QY 41 CACTTGTGGGCTGTGCGAGCCAGAGCTGTGCGGCGCGCGCGCGCGCGCGCGCG 100
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 CGCTGTGTGTGTGTGCGCGAGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
QY 101 GCTTGGGCGGAGCAGCTGTGCGGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 160
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 421 AC-----CCCAAAAGCTGACCCCTTTAGCTTCAAGAGATTATCCCAAGCTAG 474
QY 161 AGCGAGCCCTGTGGCGCGAGTGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 220
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 475 AGAAGACCTTGTGGCGCGAGCGCGCGAGATATGAGGAGAGATACAGAACTCGA 534
QY 221 TCGGAGACTGTGTGCGCACTACAGCCGAGATCTTCAAGATGAGAGAAAGATG 280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 535 TTAGGAACCTACCCCAATACAGCCGAGATATTAAGATGAGAGAAAGACTG 594
QY 281 GAGCGAGCCGCTGTATACAGGAGGCTGTGAGAGGAGGAGGAGGAGGAGGAGG 340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 595 GAGGAGCCGCTGTATACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 654
QY 341 TGATGAACATGTGGCGCGAGTGGCGAGTGTGAGAGGAGGAGGAGGAGGAGGAG 400
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 655 TGATGAACATGTGGCGCGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 714
QY 401 ACCAGCTCAGATTCATCTCAGTACAGAGCGGCTGTGAGATCAGTACGTGAGC 460
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 715 ATCATTCAGAGAGTCTCTACATATGAGGAGTGTGAGAGGAGGAGGAGGAGGAG 774
QY 461 GCGAGCGGAGAGTATGAGGAGTGTGCGCGCGCTGTGAGAGGAGGAGGAGGAG 520
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 775 GGGAGCGGAGAGTATGAGGAGTGTGCGCGCGCTGTGAGAGGAGGAGGAGGAG 834
QY 521 TCTACTAGAGTCCCGGAGACAGTGTGCGGAGTGTGAGAGGAGGAGGAGGAGGAG 580
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 835 TCTACTAGAGTCCCGGAGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 894
QY 581 TCGGCGCGCGCGCTGTGCGGAGAGTGTGCGGAGTGTGAGAGGAGGAGGAGGAG 640
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 895 CCAATCTGAGAGGCTGTGCGGAGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAG 954
QY 641 AAGGCTGCGGGAATCAGCCGCGAGACTGGTTTGGCGCGCATGCTCAGCGCGG 700
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 955 AGTAGTAGAGGAGTATGAGTGTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGG 1014
QY 701 TGTGTCGCGGCGGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 760
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1015 TGTGTACAGGACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1074
QY 761 TGTGTGTGAGACGAGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 820
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1075 ACGTGATCGAGACGCGGAGCGCGGAGGAGTGTGCGGCTAGACCGCGCGAGCT 1134
QY 821 TTGCGCGCTC-----GAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGT 868
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1135 TGTGTGGCGCGCACAGACTCGGCGCGACTCGGAGAGGAGGAGGAGGAGGAGG 1194

```



```

QY      334  ATTCCCGATGAGCAATGTGGCCCGGAGTGGCGCTAGCACTGAGTGGGCTGGACGAG 393
Db      121  ATCTCTGTCAATGAAACCAAGTGGCCCTGGTGTGAACACTGGCGGGTGAACCAAGGCCGGGATGAA 180
QY      394  GAGCGCCACCAAGCTCAGGATTCACCTCCACTACAGAAAGGCCGTGGTCTTGGACATCACTACG 453
Db      181  GATGGCCATCACTCAGAGGAGATCTTTACACTATGAGGGCCGGCGGGGATATCACACC 240
QY      454  TCTGACCGCCGACCCGACAAATATAGGGTGTGCTGGCCGCCCTCGCAGTGGAAAGCCGGCTTC 513
Db      241  TCAGACCGCTGACCAAAATAAATATGAGACTGTGTGGCCGCGCTTAGAGTGGAGGCGCGGCTTC 300
QY      514  GACTGGGCTACTACGAGATCCCGCAACAGCTCCACAGCTGACGCTGACGCTGAGAGAGAGG 573
Db      301  GACTGGGCTGTATTACGAGTCCAAAGGCCACAGCTGATTCCTGTGCAAGTGTGAGCATTCG 360
QY      574  CTGGCGGCTCCGGGGCGGGCGGCTGCTTCCGGGAAATGCAACTGTGACCTGTGTGAGAGGCG 633
Db      361  GCGCGTGGCAAGACAGAGTGGGTGGCTTCTCTGCGGAGCCCAAGTGGCGGCTTATGAGAAAGGG 420
QY      634  GAGCGGAAAGGGCTGTGCGGGAATCTCACCCGAGAGACTGGGTTTTGGCGGCGGATGCGTCA 693
Db      421  GAGCGTGTGGCCCTGTGCACTGTATTAAGCCAGAGAACCGGGGTGCTGGCCATAGGGGAGAT 480
QY      694  GGGCGGGTGGGCCAGCGCGGCTGCTGCTCTCTGTGAGACGGGAGCTTGACAGCCCGCGGCT 753
Db      481  GGGACCCCCACCTTCAATGATGATGTCTTATTTCTGTGAGACGGCGAGCCAAACCGGCTGAGA 540
QY      754  TCATTTGGGCTGTGTGAGAACCGAGATGGCGCTTCACGCAAACTGTTGTGCTCACAGCCCTGGC 813
Db      541  GCTTTCCAGGTCAATCGAGAGCTCAGAGATCTCTCGCGGTGGCTGTGGGCGCTCAAGCCTGCCAC 600
QY      814  CTGGTGTTCGCGCTCGAGGGCGCGCGCGCGCGCGGCGGCGAGCTTGGACCGGTGTTCGCG 873
Db      601  CTGGCTCTCATTTGGCGGCAATCATACAGAAACCAAGCACCCACCTTCGGGCGACATTTGGC 660
QY      874  CGCGCGGCTACCGCGCTGGGGAATCTGCTGTGGCGCGCGCGCGGGAGATGGGCTTGGCCAGCG 933
Db      661  AGCCATGTGCAACCAAGGCAATATATGTCTGTATCAGGGGATACCAAGGCGCTCCAGCGTGT 720
QY      934  CGCGGAGCCCTGTGTGGCGCGGGAGGAAGCCGTGGGCGGTGTTCCGGCGGCTCACCGGCGAC 993
Db      721  CGGGTGGCAAGCTGTCTCACCCACAGCTGGCCCTTGGGTCTATAGTCTCTCTCACAAAGGAT 780
QY      994  GGGAGCTGCTGTGTGAACGATGTCTTGCCCTCTTGTCTACGCGGTTCTGAGAGATCAACAG 105
Db      781  GGGAGACTTGTGTGAGAGATGTGTGGCTCCCTGCTTTGAGAGCTGTGGCTGACCAACAT 840
QY      1054  TGGGCGCACCGCGCTTTTGCCCTTGTAGACTGTGCAC 1092
Db      841  CTGGCTCAGTTGGGCTTCTGTGGCCCTGTGAGCTGTTCGC 879

RESULT 12
US-08-356-060A-3
: Sequence 3, Application US/08356060A
: Patent No. 5844079
: GENERAL INFORMATION:
:   APPLICANT: Ingham, Phillip W.
:   APPLICANT: McMahon, Andrew P.
:   APPLICANT: Tabin, Clifford J.
:   TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
:   TITLE OF INVENTION: Proteins and Uses Related Thereto
:   NUMBER OF SEQUENCES: 47
:   CORRESPONDENCE ADDRESS:
:   ADDRESSEE: LAHIVE & COCKFIELD
:   STREET: 60 State Street
:   CITY: Boston
:   STATE: MA
:   COUNTRY: USA
:   ZIP: 02109
:   COMPUTER READABLE FORM:
:   MEDIUM TYPE: Floppy disk

```

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,060A
FILING DATE: 14-DEC-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HHI-006CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1056 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1056
US-08-356-060A-3

Query Match          29.9%; Score 355.8; DB 3; Length 1056;
Best Local Similarity 62.8%; Pred. No. 2.2e-66;
Matches 552; Conservative 0; Mismatches 327; Indels 0; Gaps 0;

OY  214 GAGCCCTTCGGGACCTGTGCGCCAACTACAACCCGACATCATCTTCAAGGATGAGAG 273
    1 GAGCCCTTCAAGAGACTACCCCACTACATATCCGACATCATCTTCAAGAGACAGAG 60
DB  274 AACAGTGAAGCCGACCGGCTGTATGACGAGCGTTGCAGAGAGAGGATGAACGCTTTGGCC 333
    61 AACACGGGGTGGCGGACCGGCTCATGACCCAGCGGCTCAAGAGACCGTGTGAACCTACTGACC 120
OY  334 ATTGCCGTGATGAACATGTGGCCCGGAGTGGCCCTACGAGTACTGAGAGGCTGGGACGAG 393
    121 AACTCTGTGATGAACCAAGTGGCTGTGTGAACATGCGGGGTACCGAAGCGCGGATGAA 180
DB  394 GACGCGCAACAGCGTCAGAGATTCACTCCACTACGAAAGCGCGTGTGGACATCACTACG 453
    181 GATGGCCATCACTACGAGAGAGTCTTTACACTATGAGGGCGCGCGGTGATATCAACACC 240
OY  454 TCTGACCGGAGCCGCAAGATAGTGGTGTGGCGCGCCTCGCAGTAGAGAAACCGCGCTTC 513
    241 TCAGACCGTGAACGAATTAAGTATGAGTGTGTGGCGGCTTAGAGTGAAGCGCGCTTC 300
DB  514 GACTGGGTCTACTACGAGTCCCGCAACGATCCACGTGTGCGTAAAGCTGATTAAGTCA 573
    301 GACTGGGTATTAAGAGATCGAAGTCCAGAGCCCAAGTGTGCTGTAAGCTGAGATTGC 360
OY  574 CTGGGGGTCCGGGGGGGGGGGCTTTCCGGGAATGCAACGTGCGGCTGTGAGAGCGGC 633
    361 GCGCGTGCMAACAGTGGTGGCTTTCTTCCGCGGAGCCCAAGTGGCTTAGAGAACGGG 420
DB  634 GAGCGGAAAGGGCTACGGGAAGTGCACGCGGAGAGTGGTTTGGCGGCATCGTCA 693
    421 GAGCGTGGGCCCTTCAAGCTGTAAAGCAGAGACCGGGTCTGTGCCATGGGGAGAGAT 480
OY  694 GCGCGGTGTGTCACCGCGCGGTGCTCTTCTGTGACCGGGGATGCGACGCGCGGCT 753
    481 GCGACCCCACTTACGATATGCTTATTTTCTGTGAGCCCGAACCAACCGGCTGAGA 540
DB  754 TCATTGTGGCTGTGAGAGCGAGTGGGCTCCAGCAAACTTGTCTACGCGCTGGGAC 813
    541 GCTTTCAGGTATGAGAGTCAAGATCTCCGCTGGCTGGCGCTTACGATGAGCTTCCAC 600

```


TITLE OF INVENTION: Proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,427B
FILING DATE: 30-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMI-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1256 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1257
US-08-176-427B-9

Query Match 26.5%; Score 316.2; DB 2; Length 1256;
Best Local Similarity 58.4%; Pred. No. 4,1e-58;
Matches 571; Conservative 0; Mismatches 403; Indels 3; Gaps 1;

QY 112 AAGCAGCTGCGCCCTTCTACACGAAATTTGTCGCCGGCGCTCCAGAGCGGACCTTG 171
DB 109 AAGAGCTGACACCTCTGCGCTACAGACGATTCATACCTAATGCGGAGAAACCTTA 168
QY 172 GCGCGCAGTGGGGCCAGGGAGGGGGGATGCGCAAGGGGTCGCAACCGCTCCGGGAACTC 231
DB 169 GGGCGACGGGAGATACGAGGGGCAAGATTAACGGCAATTCGAGAGATTAAAGAACTT 228
QY 232 GTGCCAATCAACCCCGACATCATCTTCAAGATGAGAGACAGACGACCGCC 291
DB 229 ACTCAATTTACATCCCGACATTTCTTTAAGATGAGAGAAACGAGGAGACGAG 288
QY 292 CTGATGACCGAGCGGTGCAAGAGAGAGGTGAACGCTTGGCCATTGCCGTGATGAATG 351
DB 289 CTCATGACACAGAGATGCAAAAGACAGCTGAACCTGCGTGGCATCTCTGATTAACGAC 348
QY 352 TGGCCCGAGTGGCCCTACGAGTCTGAGGGCTGGGACGAGAGCGCCACCGCTCAG 411
DB 349 TGGCCAGGGGTTAACTGCTGTGTACAGAGGGCTGGATGAGAGAGGTCACCATTTTGA 408
QY 412 GATTCATCTCATAGAGGCGGCTTTGACATCAGTACGTGACCGGAGCGCAAC 471
DB 409 GATATCTCTCATACGAGAGGAGAGAGCTGTGATATTACACCTCTGACCGAGACAGAGC 468
QY 472 AAGTATGGTCTGCGCGCTTCACAGTGAAGCGGCTTGACTGGTCTACTACGAG 531
DB 469 AATATGGGAGACCTGTCTGCTACCTGAGAGGCTGAGTTGATGGGTCTATTACGAG 528
QY 532 TCCGCAACCAAGCTCAGTGTCTGCTCAAGGTGATTAAGTCTACTGCGGCTCCGGGCGG 591
DB 529 TCCAAAGCCACATTCATGTCTGTCAAGAGAAATTCGTTGCTGCGAAATCTGGG 588

QY 592 GCGTCCTTCCGGGAATGCAACTGTGCGCCTGTGTAGCGCGAGCGGAAAGGGCTGCGG 651
DB 589 GCGCTTTCCAGAGTTGCGCTGTGTGTGCTCCAGAGCGGAGGACAAAGGCCGTGAG 648
QY 652 GAACTGCAACCGGAGAGACTGGGTTTGGCGGCCGATGCGTCAAGCGCGGCTGCCACG 711
DB 649 GACCTGAACCCCGGAGCAAGGTGCTGGCGGACAGACCGCGGAAACCTGTGTTCAGC 708
QY 712 CCGGTGCTGCTCTCTGACGAGCGGACTTGACAGCGCGGGCTTCAATTTGTGCTGTGAG 771
DB 709 GACTTCATCATGTTTACAGACCGAGACTCCACAGACGACAGCGTGTGTTTACGTAGAA 768
QY 772 ACCGAGTGGCTTCCAGCAAACTGTTGCTACAGCGCCCTGCGCACTGTGTTGCCCTGA 831
DB 769 ACGCAAGAACCGTGAAGAAAGTACCGCTCAACCGCGCTCACTCTTTTGTCTGAC 828
QY 832 GGGCGCGCGCGCGCGCGCACTTTCACCGGTGTGCGCGCGCGGCTAAGCGCTGCG 891
DB 829 AACTCAACGGAAGATCTCCACACATGACCCCGCGTATGCCAGAGTGCAGAGCCGGA 888
QY 892 GACTGGTGTGCGCGCGCGCGGCGG---GGATGCGCTTTCGCGCAGCGCGCTGCGCTGTG 948
DB 889 CAAAGGTGATGTTGTTGATGATGATGCGGTCAAGCTTAATCTGTGATGTGAGGAGATA 948
QY 949 GCGCGGAGAGAACCGGTGCGGTGTTGCGCGCTCACCGCGCACGAGAGCGTGTGAGTG 1008
DB 949 TACACGAGAGAGACAGCGGGCTGCTCCACACAGTATGACATGAGAACCTTTGTGCT 1008
QY 1009 AAGATGCTCTGCGCTCTTGTACGCGGTTCTGAGAGTCAACCACTGGCGCACCGCT 1068
DB 1009 GACAGAAATACCTGCGCTGTTACCGCGCTTAATAGAGACACAGGGGCTTGGCATTTGCC 1068
QY 1069 TTTGCCCTCTGAGACT 1085
DB 1069 TTCGCGCGCGCGAGGCT 1085

RESULT 15

US-08-356-060A-5
Sequence 5, Application US/08356060A

Patent No. 5844079

GENERAL INFORMATION:

APPLICANT: Ingham, Phillip W.

APPLICANT: McMahon, Andrew P.

APPLICANT: Tabin, Clifford J.

TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing

NUMBER OF SEQUENCES: 47

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD

STREET: 60 State Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII(text)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/356,060A

FILING DATE: 14-DEC-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/176,427

FILING DATE: 30-DEC-1993

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.

REGISTRATION NUMBER: 36,709

REFERENCE/DOCKET NUMBER: HMI-006CP

TELECOMMUNICATION INFORMATION:

Thu Jun 8 15:53:43 2000

us-08-900-220-8.mli

THIS PAGE BLANK (USPTO)

QY	61	CAGAGCTGGGGGGCCGGGGCCGGTGTGGCGGGCGCGCTATGCGGCAAGGACGTC	120
Db	61	CAGAGCTGGGGGGCCGGGGCCGGTGTGGCGGGCGCGCTATGCGGCAAGGACGTC	120
QY	121	GTGGCGGTACTCTCAAGCAATTGTGTCGCCGGGTGCACAGGCGGACCTCGGGCGCAAT	180
Db	121	GTGGCGGTACTCTCAAGCAATTGTGTCGCCGGGTGCACAGGCGGACCTCGGGCGCAAT	180
QY	181	GGGCCAAGCGAAGGGAGGGGTGGCAGAGGGGCTCCGAGCGCTTCCGGGACTGTGCCCAAC	240
Db	181	GGGCCAAGCGAAGGGAGGGGTGGCAGAGGGGCTCCGAGCGCTTCCGGGACTGTGCCCAAC	240
QY	241	TACAAACCCCGACATCATCTTCAAGAGTGAAGAGAACAGTGGAGGCCGACCGCTGATGACC	300
Db	241	TACAAACCCCGACATCATCTTCAAGAGTGAAGAGAACAGTGGAGGCCGACCGCTGATGACC	300
QY	301	GAGCGTTGCAAGGAGAGGGGTGAACGCGTTTGGCCATTGCGGTGATGAACATGTGGCCGGA	360
Db	301	GAGCGTTGCAAGGAGAGGGGTGAACGCGTTTGGCCATTGCGGTGATGAACATGTGGCCGGA	360
QY	361	GTGGCGCTAGAGTGAAGAGGCTGGGAGAGGAGCGGCCACCGCTCAGGATTCACCTC	420
Db	361	GTGGCGCTAGAGTGAAGAGGCTGGGAGAGGAGCGGCCACCGCTCAGGATTCACCTC	420
QY	421	CACATCAAGAGGCGGTGTTGSAATCATCATAGCTGACCGGSAACCGCAACAAAGTATGG	480
Db	421	CACATCAAGAGGCGGTGTTGSAATCATCATAGCTGACCGGSAACCGCAACAAAGTATGG	480
QY	481	TTCGTGGCGGCGCTCGCAGTGGAGAGCGCGCTTGCATGGGTACTACAGCATGCCGCAAC	540
Db	481	TTCGTGGCGGCGCTCGCAGTGGAGAGCGCGCTTGCATGGGTACTACAGCATGCCGCAAC	540
QY	541	CACGTCAAGCTGTGCGTCAAAAGCTGATTACTACTGGCGGTCCGGGCGGGCGGTCTTT	600
Db	541	CACGTCAAGCTGTGCGTCAAAAGCTGATTACTACTGGCGGTCCGGGCGGGCGGTCTTT	600
QY	601	CCGGGAAATGCACTGTGGGCGCTGTGGAGCGGAGCGGAAAGCGTGGGGAACTGCAC	660
Db	601	CCGGGAAATGCACTGTGGGCGCTGTGGAGCGGAGCGGAAAGCGTGGGGAACTGCAC	660
QY	661	CGCGGAGACTGTGGTTTGGCGGGCGGATGCGTCAGCGCGGGTGGTGGCCACGCGGTGCTG	720
Db	661	CGCGGAGACTGTGGTTTGGCGGGCGGATGCGTCAGCGCGGGTGGTGGCCACGCGGTGCTG	720
QY	721	CTCTTCCTGGAACGGGACCTTGCAAGCGCGGGCTTCAATTTGGCTGTGGAGACCAAGTGG	780
Db	721	CTCTTCCTGGAACGGGACCTTGCAAGCGCGGGCTTCAATTTGGCTGTGGAGACCAAGTGG	780
QY	781	CCCTCAAGCAAACTGTGTGTCACAGCCCTGGCAACTGTGTTGGCGGTGAGAGGCGCGCG	840
Db	781	CCCTCAAGCAAACTGTGTGTCACAGCCCTGGCAACTGTGTTGGCGGTGAGAGGCGCGCG	840
QY	841	CCCGCGCGAGGCGACTTGTGCACCGGTTTCGGCGCGCGCTACGCGCTGGGACATCGGTC	900
Db	841	CCCGCGCGAGGCGACTTGTGCACCGGTTTCGGCGCGCGCTACGCGCTGGGACATCGGTC	900
QY	901	CTGGCGCGCGGGGGGATGCGCTTGGGCCACGCGCGCGTGGCCCGCTGTGTGGCGGGAGAA	960
Db	901	CTGGCGCGCGGGGGGATGCGCTTGGGCCACGCGCGCGTGGCCCGCTGTGTGGCGGGAGAA	960
QY	961	GCGCTTGTGACGGGGTGTGGAGAGTCAACAGTGGGGCACCGGGCTTTGCGCCCGTTG	1020
Db	961	GCGCTTGTGACGGGGTGTGGAGAGTCAACAGTGGGGCACCGGGCTTTGCGCCCGTTG	1020
QY	1021	GCGCTTGTGACGGGGTGTGGAGAGTCAACAGTGGGGCACCGGGCTTTGCGCCCGTTG	1080
Db	1021	GCGCTTGTGACGGGGTGTGGAGAGTCAACAGTGGGGCACCGGGCTTTGCGCCCGTTG	1080
QY	1081	AGACTGTGCACGCGGCGTAAAGGGGGGCTGTCCCGGGGGCGGTCCAGCGCACTGGCATG	1140
Db	1081	AGACTGTGCACGCGGCGTAAAGGGGGGCTGTCCCGGGGGCGGTCCAGCGCACTGGCATG	1140
QY	1141	CATTGTACTCTGCGCTCTTCAACCGCTTAAAGGAGAGCACTAGGGCTG	1190

ID		Score	% Identity	Mismatches	Gaps	Length
D8	CATTGCTACTTCGGCCTCGTCTACCGCTTAGCCAGAGACTCTGTGCATG	99.9%	100.0%	0	0	1190
D7	X25105 standard; cDNA; 1190 bp.					
D6	X25105.					
D5	05-JUL-1999 (first entry)					
D4	Human Desert hedgehog protein Dhh cDNA.					
D3	Desert hedgehog; dhh gene; human; hedgehog therapeutic;					
D2	p1c therapeutic; patched; signal transduction; muscle atrophy;					
D1	cachexia; muscular myopathy; myoblastic sarcoma; therapy; ds.					
D0	Homo sapiens.					
PX	W09910004.1					
PD	04-MAR-1998					
PR	26-AUG-1998 J017922.					
PS	29-AUG-1997 US-057394.					
PI	(ONTO.) ONTOGENY INC.					
PL	Bladgen CS, Currie PD, Hughes SM, Ingham PW;					
DR	wpi; 99-243557/20.					
PT	P-PDB; Y05517.					
PF	A new method to regulate muscle growth					
PG	Disclosure: Page 106-108; 130pp: English.					
PC	This nucleotide sequence comprises a coding region for the human					
CB	Desert hedgehog protein Dhh (see X05517). The invention relates to					
CA	a method for modulating the formation and/or maintenance of muscle					
CC	tissue by ectopically contacting muscle cells, especially muscle					
CD	stem/progenitor cells, in vitro or in vivo, with a hedgehog					
CE	therapeutic (i.e. hedgehog polypeptides and gene therapy					
CF	constructs) or ptc therapeutic (i.e. a small organic molecule that					
CG	mimics the effect of hedgehog proteins on patched signalling, or					
CH	activates or potentiates patched signalling) in an amount effective					
CI	to alter the growth state of the treated cells. Also claimed is a					
CJ	method for treatment or prevention of disorders of, or surgical or					
CK	cosmetic repair of, such muscle tissues, by administering a					
CL	hedgehog polipeptide or ptc therapeutic. The disorder may be					
CM	muscle atrophy, in particular skeletal muscle atrophy or cardiac					
CN	muscle atrophy, cachexia, or muscular myopathy (all claimed). The					
CO	hedgehog polipeptide or ptc therapeutic can inhibit growth of					
CP	myoblastic-derived tissue to provide treatment of hyperblastic or					
CQ	neoplastic growth of muscle tissue such as in myoblastic sarcoma					
CR	(also claimed). The hedgehog therapeutic preferably comprises at					
CS	least a bioactive extracellular portion of a hedgehog protein (see					
CT	X05510-19) encoded by a vertebrate hedgehog gene (see X25098-107),					
CU	especially a human hedgehog gene.					
CV	Sequence 1190 BP; 176 A; 375 C; 424 G; 215 T;					
Query Match	99.9%; Score 1190; DB 1; Length 1190;					
Best Local Similarity	100.0%; Pred. No. 1, de-238;					
Matches 1190:	Conservative 0; Mismatches 0; Indels 0; Gaps					

```

QY 301 GAGGCTTGCAGAGAGAGGCTGACGCTTTGGCCATTTGCCGTGATGACATGTGCCCGGA 360
DB 301 GAGGCTTGCAGAGAGAGGCTGACGCTTTGGCCATTTGCCGTGATGACATGTGCCCGGA 360
QY 361 GTGGCCCTACAGAGTACTGAGAGGCTGGGACGAGAGGCGCACACGCGTCAGATTCACTC 420
DB 361 GTGGCCCTACAGAGTACTGAGAGGCTGGGACGAGAGGCGCACACGCGTCAGATTCACTC 420
QY 421 CACTACGAAAGGCGGTGCTTTGGACATCACTAGTGTGACCGCGACCGGACAAAGATGGG 480
DB 421 CACTACGAAAGGCGGTGCTTTGGACATCACTAGTGTGACCGCGACCGGACAAAGATGGG 480
QY 481 TTGCTGGCGCCCTTCGACAGTGAAGCCGGCTTCGACTGGGTCTACTACGATCCCGAAC 540
DB 481 TTGCTGGCGCCCTTCGACAGTGAAGCCGGCTTCGACTGGGTCTACTACGATCCCGAAC 540
QY 541 CACGCTCAGTGTGGGTAAAGCTGATTAACACAGGCGGCTCCGGGGGGGGGGGCTGCTT 600
DB 541 CACGCTCAGTGTGGGTAAAGCTGATTAACACAGGCGGCTCCGGGGGGGGGGGCTGCTT 600
QY 601 CCGGGAATGCACTGTGCTGCGCTGTGAGCGCGGACGAGGAGGAGGCTGCGGAACTGCAC 660
DB 601 CCGGGAATGCACTGTGCTGCGCTGTGAGCGCGGACGAGGAGGAGGCTGCGGAACTGCAC 660
QY 661 CCGGAGAGTGGGTGTTGGCGCGCGAGTGTGACGCGGCTGAGCGCGGCTGCTGCTG 720
DB 661 CCGGAGAGTGGGTGTTGGCGCGCGAGTGTGACGCGGCTGAGCGCGGCTGCTGCTG 720
QY 721 CTCTCCGAGCGGAGCTTGCAGAGCGCGGCTTCACTTGTGGCTGTGAGAACCGAGTGG 780
DB 721 CTCTCCGAGCGGAGCTTGCAGAGCGCGGCTTCACTTGTGGCTGTGAGAACCGAGTGG 780
QY 781 CCTCCAGCAAACTGTGCTCAGCGCTGAGCACTGTGTGTGCGCTGAGAGCGCGGCG 840
DB 781 CCTCCAGCAAACTGTGCTCAGCGCTGAGCACTGTGTGTGCGCTGAGAGCGCGGCG 840
QY 841 CCGCGGCGGAGGCTTGTGACCGGCTGTGCGCGCGCGCTACGCGCTGGGAGCTCGGTG 900
DB 841 CCGCGGCGGAGGCTTGTGACCGGCTGTGCGCGCGCGCTACGCGCTGGGAGCTCGGTG 900
QY 901 CTGGCGCGCGGAGGCTTGTGCGCGCGCGCTGAGCGCGGCTGAGCGCGGAGGAGAA 960
DB 901 CTGGCGCGCGGAGGCTTGTGCGCGCGCGCTGAGCGCGGCTGAGCGCGGAGGAGAA 960
QY 961 GCGCTGGCGGCTGTGCGCGCGCTGACCGCGGAGGAGGAGCTGTGTGAAGATGTCTG 1020
DB 961 GCGCTGGCGGCTGTGCGCGCGCTGACCGCGGAGGAGGAGCTGTGTGAAGATGTCTG 1020
QY 1021 GCGCTGTGCTACGCGGCTGTGAGAGTACCAAGTGGGCGACCGCGCTTTGGCCCTTG 1080
DB 1021 GCGCTGTGCTACGCGGCTGTGAGAGTACCAAGTGGGCGACCGCGCTTTGGCCCTTG 1080
QY 1081 AGACTGGCGAGGCGCTAGGGGCGTGTCCCGCGGGGCGCTTCACCGCACTGGCATG 1140
DB 1081 AGACTGGCGAGGCGCTAGGGGCGTGTCCCGCGGGGCGCTTCACCGCACTGGCATG 1140
QY 1141 CATGTGACTCTCGGCTCTTACCGGCTTACGAGAGAGTACTGGGCTG 1190
DB 1141 CATGTGACTCTCGGCTCTTACCGGCTTACGAGAGAGTACTGGGCTG 1190

```

```

RESULT 3
ID V62395 standard: cDNA: 1188 BP.
AC V62395;
DT 02-FEB-1999 (first entry)
DE Human Desert hedgehog protein cDNA.
KW Desert hedgehog; HUH1; human; ds.
OS Homo sapiens.
FH Key location/Qualifiers
FT sig_peptide 1..66
FT mat_peptide 67..594

```

```

FT EP-824048-A2 /tag- b
PN 28-OCT-1998
PR 24-APR-1998 303187.
PR 14-APR-1998 JP-117873.
PR 25-APR-1997 JP-121578.
PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
PI Ariyasu T, Nakamura S, Orita K;
DR WPJ: 98-544642/47.
DR P-PSDB: W79595.
PT Human Desert hedgehog protein - and corresponding DNA and monoclonal
PT antibody
PS Claim 10; Page 24-25; 39pp; English.
CC This cDNA sequence codes for a precursor (see W79595) of a novel
CC human Desert hedgehog protein (see also W79593). Human Desert
CC hedgehog protein DNA (HUH1) was originally identified following
CC screenings of human cell lines by PCR using primers (see V62402-03)
CC based on the mouse Desert hedgehog gene. These screenings
CC indicated that the ARH-77 (ATCC CRL-1621) cell line, which is
CC derived from the plasma cell of a leukemia patient, expressed a
CC specific gene at an elevated level. Sequencing confirmed it to
CC be a novel human gene showing homology to the mouse Desert hedgehog
CC gene. DNA encoding N-terminal sequences (see V62397) was obtained
CC by further PCR amplifications (see V62404-06). The invention
CC provides Desert hedgehog polynucleotides and protein, a monoclonal
CC antibody (Mab) that recognises the protein, a process for producing
CC the protein, and a method for detecting the protein using the Mab.
CC The hedgehog protein, DNA and Mab can be used to elucidate
CC hereditary morphological abnormalities in humans to establish their
CC treatments and diagnoses.
SQ Sequence 1188 BP; 179 A; 376 C; 419 G; 214 T;

Query Match 98.8%; Score 1176.8; DB 1; Length 1188;
Best Local Similarity 99.4%; Pred. No. 7.7e-236;
Matches 1181; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATGGCTCTCCAGCAAACTGTGCTCAGCGCTGAGCACTGTGTGTGCGCTGAGAGCGCGGCG 60
DB 1 ATGGCTCTCCAGCAAACTGTGCTCAGCGCTGAGCACTGTGTGTGCGCTGAGAGCGCGGCG 60
QY 61 CAGAGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
DB 61 CAGAGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
QY 121 GTGGCGGCTACTCTACAGCAATTTGTGCGCGCGCTGCGCAAGCGGAGCTTGGCGGCGG 180
DB 121 GTGGCGGCTACTCTACAGCAATTTGTGCGCGCGCTGCGCAAGCGGAGCTTGGCGGCGG 180
QY 181 GGGCCAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
DB 181 GGGCCAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
QY 241 TACAACCCCGCATCTTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 241 TACAACCCCGCATCTTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 301 GAGGCTTGCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
DB 301 GAGGCTTGCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
QY 361 GTGGCGCTACAGAGTACTGAGAGGCTGGGACGAGAGGCGCACACGCGTCAGATTCACTC 420
DB 361 GTGGCGCTACAGAGTACTGAGAGGCTGGGACGAGAGGCGCACACGCGTCAGATTCACTC 420
QY 421 CACTACGAAAGGCGGTGCTTTGGACATCACTAGTGTGACCGCGACCGGACAAAGATGGG 480
DB 421 CACTACGAAAGGCGGTGCTTTGGACATCACTAGTGTGACCGCGACCGGACAAAGATGGG 480
QY 481 TTGCTGGCGCGCTTCGACAGTGAAGCCGGCTTCGACTGGGTCTACTACGATCCCGAAC 540
DB 481 TTGCTGGCGCGCTTCGACAGTGAAGCCGGCTTCGACTGGGTCTACTACGATCCCGAAC 540

```

```

QY 541 CAGCTCAGCTGTCGGTCAAACTGATACCTACTGCGGGTCCGGGCGGGCGCTCTT 600
DB 541 CAGATCAGCTGTCGGTCAAACTGATACCTACTGCGGGTCCGGGCGGGCGCTCTT 600
QY 601 CCGGAAATGCACTGTGCGCTGTGAGCGGCGGAGGAAAGGGTGTGGGAAATGCA 650
DB 601 CCGGAAATGCACTGTGCGCTGTGAGCGGCGGAGGAAAGGGTGTGGGAAATGCA 650
QY 661 CCGGAGACTGGGTTTGGCGGCGGATGCTACAGCGCGGGTGGTCCACCGCGGTCT 720
DB 661 CCGGAGACTGGGTTTGGCGGCGGATGCTACAGCGCGGGTGGTCCACCGCGGTCT 720
QY 721 CTCTTCTGAGACCGGACTTGCAGCGCGGGCTTCAATTGTGGCTGTGGAGACGAGT 780
DB 721 CTCTTCTGAGACCGGACTTGCAGCGCGGGCTTCAATTGTGGCTGTGGAGACGAGT 780
QY 781 CTTCAAGCAAACTGTGCTACAGCGCTGACCTGGTGTTCGCGTCAAGGCGCGG 840
DB 781 CTTCAAGCAAACTGTGCTACAGCGCTGACCTGGTGTTCGCGTCAAGGCGCGG 840
QY 841 CCGCGCCAGAGGACTTGTGACCGGATGTCGCGCGCGCTACGCGCTGGGACTGGT 900
DB 841 CCGCGCCAGAGGACTTGTGACCGGATGTCGCGCGCGCTACGCGCTGGGACTGGT 900
QY 901 CTGGCGCCGCGGCGGATGCTGCGCGCTGCGCGCGCTGCGCGCTGCGCGCGG 960
DB 901 CTGGCGCCGCGGCGGATGCTGCGCGCTGCGCGCGCTGCGCGCTGCGCGCGG 960
QY 961 GCGCGCGCGCTGCTGCGCGCTGCGCGCTGCGCGCGCTGCGCGCTGCGCGCGG 1020
DB 961 GCGCGCGCGCTGCTGCGCGCTGCGCGCTGCGCGCGCTGCGCGCTGCGCGCGG 1020
QY 1021 GCGCTTGTGCTACCGGCTTGTGAGAGTACCAAGTGGGCGGCGGCTTGTGCG 1080
DB 1021 GCGCTTGTGCTACCGGCTTGTGAGAGTACCAAGTGGGCGGCGGCTTGTGCG 1080
QY 1081 AGACTGCTGACGCGGCTGAGGCGGCTGCTGCGCGCGGCGGCTGCGCGGCTG 1140
DB 1081 AGACTGCTGACGCGGCTGAGGCGGCTGCTGCGCGCGGCGGCTGCGCGGCTG 1140
QY 1141 CATTGGTACTCTGCGCTCTCTACCGCTTACGCGAGAGGACTTGGG 1188
DB 1141 CATTGGTACTCTGCGCTCTCTACCGCTTACGCGAGAGGACTTGGG 1188
RESULT 4
V62394
ID V62394 standard; cDNA: 1122 BP.
AC V62394:
DE 02-FEB-1999 (first entry)
DE Human Desert hedgehog protein cDNA.
KW Human Desert hedgehog; HuhdH; human; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT mat_peptide 1..528
FT tag a
PE 22-APR-1998: 303187.
PD 28-OCT-1998:
PR 14-APR-1998: JP-117873.
PR 25-APR-1997: JP-121578.
PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
PI Arizasu T, Nakamura S, Orita K.
DR WPI: 98-544642/47.
DR P-PSDB: W79594.
PT Human Desert hedgehog protein - and corresponding DNA and monoclonal
PT antibody.
PS Claim 9, page 22-23; 39pp; English.
CC This cDNA sequence codes for a precursor (see W79594) of a novel
CC human Desert hedgehog protein (see also W79593). Human Desert
CC hedgehog protein DNA (HuhdH) was originally identified following
CC screenings of human cell lines by PCR using primers (see V62402-03)
CC based on the mouse Desert hedgehog gene. These screenings

```

```

CC indicated that the ARH-77 (ATCC CRL-1621) cell line, which is
CC derived from the plasma cell of a leukemia patient, expressed a
CC specific gene at an elevated level. Sequencing confirmed it to
CC be a novel human gene showing homology to the mouse Desert hedgehog
CC gene. DNA encoding C-terminal sequences (see V62398) was obtained
CC by further PCR amplifications (see V62407-08). The invention
CC provides Desert hedgehog polynucleotides and protein, a monoclonal
CC antibody (MAb) that recognizes the protein, a process for producing
CC the protein, and a method for detecting the protein using the MAb.
CC The hedgehog protein, DNA and MAb can be used to elucidate
CC hereditary morphological abnormalities in humans to establish their
CC treatments and diagnoses.
SQ Sequence 1122 BP; 170 A; 351 C; 403 G; 198 T;

Query Match
Best local similarity 93.4%; Score 1122.4; DB 1; Length 1122;
Matches 1116; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 67 TCGGGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 126
DB 1 TCGGGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
QY 127 CTACTCTAAGCAATTTTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 186
DB 61 CTACTCTAAGCAATTTTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
QY 187 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 246
DB 121 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
QY 247 CCGGACATATTTTCAAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 306
DB 181 CCGGACATATTTTCAAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
QY 307 TCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 366
DB 241 TGTAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
QY 367 CTACGAGTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 426
DB 301 CTACGAGTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
QY 427 GAGGCGGCTGCTTGTGACATCACTACGCTGTGACCGGAGCGGACAGAGTATGGT 486
DB 361 GAGGCGGCTGCTTGTGACATCACTACGCTGTGACCGGAGCGGACAGAGTATGGT 420
QY 487 GCGGCGCTGCGAGTGAAGCGGCGGCTGCACTGGGCTCTACGAGTCCCGCAACGAT 546
DB 421 GCGGCGCTGCGAGTGAAGCGGCGGCTGCACTGGGCTCTACGAGTCCCGCAACGAT 480
QY 547 CAGGTGTCGCTCAAGGAGTATACCTACGCGGCTGCGGCGGCGGCGGCTTGTCCGGA 606
DB 481 CAGGTGTCGCTCAAGGAGTATACCTACGCGGCTGCGGCGGCGGCGGCTTGTCCGGA 540
QY 607 AATGCACTGTGCGCGCTGTGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 666
DB 541 AATGCACTGTGCGCGCTGTGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
QY 667 GACTGGGTTTGGCGGCGGAGTGCATGAGGCGGCGGCGGCGGCGGCGGCGGCGG 726
DB 601 GACTGGGTTTGGCGGCGGAGTGCATGAGGCGGCGGCGGCGGCGGCGGCGGCGG 660
QY 727 CTGAGACCGGAGTTCAGGCGCGGCTTCAATTGTGCGCTGTGAGACGAGAGGCGCTCA 786
DB 661 CTGAGACCGGAGTTCAGGCGCGGCTTCAATTGTGCGCTGTGAGACGAGAGGCGCTCA 720
QY 787 GCGAAACTGTGCTCAAGCGCTGCGACCTGTGTTGCGCGCTGAGAGGCGCGCGCGCG 846
DB 721 GCGAAACTGTGCTCAAGCGCGCTGCGACCTGTGTTGCGCGCTGAGAGGCGCGCGCG 780
QY 847 CAGGCGAGTTTGACCGGCTGCGCGCGGCGGCTACGCGCGTGGGAGTGGTGTGGCG 906
DB 847 CAGGCGAGTTTGACCGGCTGCGCGCGGCGGCTACGCGCGTGGGAGTGGTGTGGCG 906

```

Db 781 CCAGCGACTTGCACCGGCTTCCGCGCCGCGCTACGCGCTGGGAGCTCGGCTGCGG 840
Qy 907 CCCGCGGAGATGCGCTTGGCCAGCGCGCGCTGCGCTGCTGGCGGAGAGAGCCGTG 966
Db 841 CCCGCGGAGATGCGCTTGGCCAGCGCGCGCTGCGCTGCTGGCGGAGAGAGCCGTG 900
Qy 967 GCGGCTGCTGCGCGCTGCGCTACCGCGGAGCGCTGCTGCTGAGATGCTCGGCGCTG 1026
Db 901 GCGGCTGCTGCGCGCTGCGCTACCGCGGAGCGCTGCTGCTGAGATGCTCGGCGCTG 960
Qy 1027 TCGTACGCGCTTGTGAGAGTACAGTGGGCGGCGGCGCTTTCCTCCCTTGAACATG 1086
Db 961 TCGTACGCGCTTGTGAGAGTACAGTGGGCGGCGGCGCTTTCCTCCCTTGAACATG 1020
Qy 1087 CTGACGCGCTGAGGCGGCTGCTCCCGCGGCGGCGGCTGCTGCGGAGCGGAGCATGATG 1146
Db 1021 CTGACGCGCTGAGGCGGCTGCTCCCGCGGCGGCGGCTGCTGCGGAGCGGAGCATGATG 1080
Qy 1147 TACTCTGCGCTCTCTACCGCTTACGCGGAGGAGGAGCTACTGCGC 1188
Db 1081 TACTCTGCGCTCTCTACCGCTTACGCGGAGGAGGAGCTACTGCGC 1122

RESULT 5
X25099
ID X25099 standard; cDNA; 1191 BP.
AC X25099;
DT 05-JUL-1999 (first entry)
DE Mouse Desert hedgehog protein Dhh cDNA.
KW Desert hedgehog; Dhh gene; mouse; hedgehog therapeutic;
KW ptc therapeutic; patched; signal transduction; muscle atrophy;
KW cachexia; muscular myopathy; myoblastic sarcoma; therapy; ds.
OS Mus sp.
BN M09910067-A2
PD 04-MAR-1998
PF 28-AUG-1998--017922.
PR 29-AUG-1997; US-057394.
PA (ONTO-) ONTOGENY INC.
PI Bladen CS, Currie PD, Hughes SM, Ingham PW;
DR WPI: 99-243557/20.
DR P-PDB: Y05511.
PT A new method to regulate muscle growth
PS Disclosure, Page 94-95; 130pp; English.
CC This nucleotide sequence comprises a coding region for the mouse
CC Desert hedgehog protein Dhh (see Y05511). The invention relates to
CC a method for modulating the formation and/or maintenance of muscle
CC tissue by ecotopically contacting muscle cells, especially muscle
CC stem/progenitor cells, in vitro or in vivo, with a hedgehog
CC therapeutic (i.e. hedgehog polypeptides and gene therapy
CC constructs) or ptc therapeutic (i.e. a small organic molecule that
CC mimics the effect of hedgehog proteins on patched signaling, or
CC activates or potentiates patched signaling) in an amount effective
CC to alter the growth state of the treated cells. Also claimed is a
CC method for treatment or prevention of disorders of, or surgical or
CC cosmetic repair of, such muscle tissues, by administering a
CC hedgehog polypeptide or ptc therapeutic. The disorder may be
CC muscle atrophy, in particular skeletal muscle atrophy or cardiac
CC muscle atrophy, cachexia, or muscular myopathy (all claimed). The
CC hedgehog polypeptide or ptc therapeutic can inhibit growth of
CC myoblastic-derived tissue to provide treatment of hyperplastic or
CC neoplastic growth of muscle tissue such as in myoblastic sarcoma
CC (also claimed). The hedgehog therapeutic preferably comprises at
CC least a bioactive extracellular portion of a hedgehog protein (see
CC Y05510-19) encoded by a vertebrate hedgehog gene (see X25098-107),
CC especially a human hedgehog gene.
SQ Sequence 1191 BP; 195 A; 371 C; 399 G; 226 T;

Query Match 81.3%; Score 968.6; DB 1; Length 1191;
Best Local Similarity 88.3%; Pred. NO. 1.1e-192;
Matches 1052; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

Qy 1 ATGCGCTCTCCGACCAATCTACTGCTGCTGCTGCGACTTTCGCGCTGCGAGCC 60

Db 1 ATGCGCTCTCCGACCAATCTACTGCTGCTGCTGCGACTTTCGCGCTGCGAGCC 60
Qy 61 CAGAGCTGCGGCGCGGCGCGGCGGCGGCTTGGCGCGGCGCGCTATGCGCGCAAGAGCTC 120
Db 61 CAGAGCTGCGGCGCGGCGCGGCGGCGGCGGCTTGGCGCGGCGCGCTATGCGCGCAAGACTT 120
Qy 121 GTGCGCTACTCTACCAAGCAATTTGCGCGCGCGCTGCGCGAGCGGAGCCCTGGGCGCACT 180
Db 121 GTGCGCTACTCTACCAAGCAATTTGCGCGCGCGCTGCGCGAGCGGAGCCCTGGGCGGAGT 180
Qy 181 GGGCGACGCGGAGGAGGAGGCTGCGAGGAGGCGCTTCCGAGGAGCTGCTGCGGCAAC 240
Db 181 GGGCGACGCGGAGGAGGAGGCTGCGAGGAGGCGCTTCCGAGGAGCTGCTGCGGCAAC 240
Qy 241 TACAACCCGACATCATCTCAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
Db 241 TACAACCCGACATCATCTCAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
Qy 301 GAGCGTTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
Db 301 GAGCGTTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
Qy 361 GTGCGCTTACAGTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
Db 361 GTGCGCTTACAGTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
Qy 421 CACTACGAGGCGGCTTGGAGATCACTAGCTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
Db 421 CACTACGAGGCGGCTTGGAGATCACTAGCTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
Qy 481 TTGCTGCGCGCGCTCGACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
Db 481 TTGCTGCGCGCGCTCGACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
Qy 541 CAGCTCACGCTGTGCTCAAAAGCTGATTAATCACTGCGGCGGCGGCGGCGGCGGCGGCGGCTT 600
Db 541 CAGCTCACGCTGTGCTCAAAAGCTGATTAATCACTGCGGCGGCGGCGGCGGCGGCGGCGGCTT 600
Qy 601 CCGGGAAGTCAACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
Db 601 CCGGGAAGTCAACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
Qy 661 CCGGAGACTGCGGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 720
Db 661 CCGGAGACTGCGGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 720
Qy 721 CTCTCTCTGAGCGGAGCTTGCAGCGCGGCGGCTTATTTGCGCTGAGACCGAGTGG 780
Db 721 CTCTCTCTGAGCGGAGCTTGCAGCGCGGCGGCTTATTTGCGCTGAGACCGAGTGG 780
Qy 781 CCGGCAAGCAATGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
Db 781 CCGGCAAGCAATGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
Qy 841 CCGGCGGAGGAGCTTGCAGCGGCTTGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 900
Db 841 CCGGCGGAGGAGCTTGCAGCGGCTTGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 900
Qy 901 CTGCGCGCGGCGGCGGAGTGGCTTGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 960
Db 901 CTGCGCGCGGCGGCGGAGTGGCTTGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 960
Qy 961 GCGGCTGCGGCTTGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 1020
Db 961 GCGGCTGCGGCTTGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 1020
Qy 1021 GCGCTTGTACAGGAGTGTGAGAGTCAACAGTGGGCGGCGGCGGCGGCGGCGGCGGCTT 1080
Db 1021 GCGCTTGTACAGGAGTGTGAGAGTCAACAGTGGGCGGCGGCGGCGGCGGCGGCGGCTT 1080
Qy 1081 AGACTGCTGACGCGCTAGGAGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 1140
Db 1081 AGACTGCTGACGCGCTAGGAGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 1140

ID	Accession	Standard	CDNA	BP
DB	1081	CGGCTGCTCAGACGGCGTCCGGGCTGTCGTCCTCGGGGCGTGCAGTCCAGCCGACTGGCAGT	1140	
QY	1141	CATTGGTACTCTCGGCTCCTCTACCGCTTACGGGAGGAGACTTACTGGGCTGA	1191	
DB	1141	CATTGGTACTCTCGGCTCCTCTTACCGCTTACCGGAGGAGACTTATGGGCTGA	1191	
RESULT 6				
ID	091642	standard	CDNA	1190 BP
AC	091642			
DE	18-MAR-1996	(first entry)		
DE	Mouse	desert hedgehog protein gene.		
KM	nervous system disorder	gene therapy; antibody; ds.		
OS	Mus musculus.			
Key		Location/Qualifiers		
FT	exon	1..303		
FT	exon	/*tag= a		
FT	exon	304..567		
FT	exon	/*tag= b		
FT	exon	568..1188		
FT	exon	/*tag= c		
PN	W0918856-A1.			
PD	13-JUL-1995.			
PR	30-DEC-1994.	U14992.		
PR	30-DEC-1993.	U5176427.		
PA	14-DEC-1994.	U5-356060.		
PA	(HARD)	HARVARD COLLEGE.		
PA	(IMCR)	IMPERIAL CANCER RES TECHNOLOGY.		
PI	Ingham PW, McMahon AP, Tabin CJ.			
PI	WPI: 95-255060/33.			
DR	P-PDB: R17345.			
DR	Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful			
PT	to treat degenerative nervous system disorder(s) and in gene			
PT	therapy.			
PS	Claim 5: Page 135-37. 210pp. English.			
CC	The sequence encodes a mouse desert hedgehog protein, homologous			
CC	to a Drosophila hedgehog protein (R7737), and has been isolated by			
CC	low stringency screening of a mouse genome DNA library, in phage			
CC	lambda. The sequence contains 3 homologous regions, encoding a			
CC	single open reading frame interrupted by introns. Splicing has			
CC	been confirmed by polymerase chain reaction amplification of first			
CC	strand cDNA generated from adult testicle RNA. Probes and primers			
CC	derived from hedgehog sequences may be used as diagnostic agents			
CC	for neuromuscular, autonomic or central nervous system disorders,			
CC	and the gene may also be used in gene therapy. Antibodies			
CC	generated from the encoded protein may be used as therapeutic or			
CC	research reagents.			
SO	Sequence	1190 BP; 194 A; 371 C; 399 G; 226 T;		
Query Match 81.2%; Score 967.6; DB 1; Length 1190;				
Best Local Similarity 88.3%; Pred. No. 1.8e-192;				
Matches 1051; Conservative 0; Mismatches 139; Indels 0; Gaps 0.				
QY	1	ATGCTCTCTCTGACCAATCTACTGCTCCCTTGTCTGCTTGGCACTTCTGGCGCCAGCC	60	
DB	1	ATGCTCTCTCTGACCAATCTACTGCTCCCTTGTCTGCTTGGCACTTCTGGCGCCAGCC	60	
QY	61	CAGACTGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGG	120	
DB	61	CAGACTGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGGCGGGG	120	
QY	121	GTGCGCTCTCTTACCAAGCAATTTGTGGCGGGCGGTGCGGAGCGGAGCGGCGGCGGCT	180	
DB	121	GTGCGCTCTCTTACCAAGCAATTTGTGGCGGGCGGTGCGGAGCGGAGCGGCGGCGGCT	180	
QY	181	GGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG	240	
DB	181	GGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG	240	
QY	241	TACAAACCCGACATCATCTTCAAGATGAGAGAACATGAGACCGGCGGCTGATGACC	300	

Db	241	TTCAACCCCGCATAACTTCTCAAGGATGAGAGAAACAGCGCGCCAGACCCGCTTATGCA	300
Qy	301	GAGCGTTCGAAGAGAGAGGTTGAACGCTTTGGCCATTGCCGTATGAAACTGTGGCCCGGA	360
Db	301	GAGCGTTCGAAGAGAGCGGTGAACGCGCTCTAGCCATCGCGGTGATGAACATGTGGCCCGGA	360
Qy	361	GTGGCGCCTACGATGACTGAGGGCTGGGAGACGAGAGACCGCCACCAGCTTCAGATTCACTC	420
Db	361	GTACGCCCTACCTGTGACTGAGAGGCTGGGAGAGAGAGGCGCACACGACACAGGATTCACCTC	420
Qy	421	CACCTACGAAGGCCGTGCTTGGATCGATCACTACGTCTGACCGCGAGACCGCAACAGTATGGG	480
Db	421	CACCTACGAAGGCCCTGCTTGGATCGATCACTACGATCAGCAGTCTGACCGGTATATAGTATGTG	480
Qy	481	TTGCTGGGCGCCCTCGGAGTGAAGACCGCGTTTGACTGGGTCTACTACAGTCCCGCAAC	540
Db	481	TTGTTGGGCGCCCTACGCTGTGTAAGAGCGGATGTGACTGGGTCTACTACAGATCCCGCAAC	540
Qy	541	CACGTCACAGTGTGGTGTCAAAAGCTGATTACTACTGTGGCGGTCCGGGCGGGCGGTCTTT	600
Db	541	CACATCCACAGTATGGTGTCAAAAGCTGATTACTACTGTGGCGGTCCGAGCCGAGGCGCTTT	600
Qy	601	CCGGGAATGTGAACCTGTGGCGCTGTGGAGCGGCGAGCGGAAGAGGGCTGTGGGAATGTGAC	660
Db	601	CCGGGAATGTGCACAGCTGTGGCGCTTTCGGAGCGGCGGAACGGAAGGGCTGTAGGGAATTACAT	660
Qy	661	CGCGAGACATGGGTTTTGGGCGCGGATGCTCAGCGCGGGGTGTGGCCACGCGCGGTCTG	720
Db	661	CGTGTGACTGGTACGTGACGTGGCGCGGTGATGCAGCGGGCGGAGTGTATCCACGCGCACTGCTG	720
Qy	721	CTCTTCTTGACCGGGACCTTGACAGCGCGGGGCTTCAATTGTGGCTGTGGAGACGAGTGG	780
Db	721	CTCTTCTTGAGACCGGGATCTGACAGCGCGGGGCTCTTGTGTGGCTGTGGAGACGAGCGG	780
Qy	781	CGTCACGCAAACTGTGGTCAAGCGCCCTGGGACCTGTGTTGGCGGTCTGAGGGCGCGGCG	840
Db	781	CTTCGCGCAAACTGTGTCTACACCTTGCACTGTGTGTGCTGTGGCGGGCGACGCG	840
Qy	841	CCCGCGCCAGCGCATTTGTGACACCGGTTTGGCGCGCGGCTACAGCGGTGGGACTCGGCTG	900
Db	841	CGTGCTCCAGGTGACTTTGACACCGGTTTGGCGCGCGGCTTACGTGTGTGGGACTCGGCTG	900
Qy	901	CTGGCGCCCGCGGGGATGCGCTTTCCGCCACAGCGCGCGCTGTGGCCCGCGGAGGAA	960
Db	901	CTGGCTCCCGCGGGGACCGGCTTCCAGCGCGCGCGCGTGTGGCCCGCGGAGGAA	960
Qy	961	GCCGCGGGCGGTTCGGCGCGGCTCACCGGCAACGGGAGCCTGTGTAAAGATGTCTCTG	1020
Db	961	GCCGCGGGCGGTTCGGACCGCTCACCGGCAAGGGAGCCTGTGTAAACAGACTCTCTC	1020
Qy	1021	GCCCTTTGCTACGCGGTTTGTGAGAGTCAACAGATGGGCGCACCGGCTTTGCCCTTTG	1080
Db	1021	GCCCTCTGCTACGCGGTTTGTAGAGATCAACAGATGGGCGCACCGGCTTGGCCCTTTG	1080
Qy	1081	AGACTGTGACAGCGGCTAGAGGGCGCTGCTCCCGCGGGGCGCTCAGCGCATGTGGCATG	1140
Db	1081	CGGCGTGTGACAGCGGCTAGGGGCTGTGCTCCCTGGGGGTGTGACAGCGCATGTGGCATG	1140
Qy	1141	CATTGTACTCTCGGCTCTCTACAGCGTTAAGCGGAGAGACTACTGGGCTG	1190
Db	1141	CATTGTACTCTCGCTCTCTTACCGCTTGGCGGAGAGATTAAATGGGCTG	1190
RESULT 7			
ID	X16183		
AC	X16183 standard; cDNA, 1190 BP.		
DT	29-APR-1999 (first entry)		
DE	Mouse Dhh hedgehog cDNA sequence.		
KW	Patched, hedgehog; ptc therapeutic; neuroprotective; neuronal cell;		
KW	brain infarction; cerebral infarction; transient ischemic attack;		
KW	stroke; cerebral infarct volume; spinal cord; oedema; trauma;		

PT inhibitor, or an agent derived from hedgehog polypeptides, useful in
PS the treatment of Parkinson's disease.
PS Disclosure: Page 85-86; 138pp; English.
CC This nucleotide sequence comprises a coding region for the mouse
CC Dhh Desert hedgehog protein (see W97766). The invention is based on
CC the finding that hedgehog proteins are useful as protective agents
CC in the treatment and prophylaxis of neurodegenerative disorders
CC resulting from the loss of dopaminergic and/or GABA-ergic neurons,
CC or the general loss of tissue from the substantia nigra.
CC Exemplary disorders include Parkinson's disease, Huntington's
CC disease (both claimed), amyotrophic lateral sclerosis and cerebral
CC ischaemia. The invention relates to hedgehog therapeutics (i.e.
CC encoding polypeptides and gene therapy constructs e.g. constructs
CC encoding recombinant hedgehog polypeptides and trans-activation
CC constructs for altering hedgehog gene regulatory sequences) and
CC ptc therapeutics (i.e. agents which mimic the effect of naturally
CC occurring hedgehog proteins on patched signalling) that are
CC effective in both human and animal subjects. Human Ihh and Dhh
CC polypeptides (see W97766-64) are preferred. The products can also
CC be used for the maintenance of differentiated neurons in cultures,
CC and to enhance the implantation of such neuronal cells in an
CC animal. They can also be used to prevent or treat neurodegenerative
CC conditions arising from the use of certain drugs, and in the
CC prevention and/or treatment of hypoxia, e.g. as a neuroprotective
CC agent.
SQ Sequence 1190 BP; 194 A; 371 C; 399 G; 226 T;

Query Match	81.2%;	Score 967.6;	DB 1;	Length 1190;
Best Local Similarity	88.3%;	Pred. No. 1.8e-192;		
Matches 1051;	Conservative 0;	Mismatches 139;	Indels 0;	Gaps 0;

QY	1	ATGGCTTCCTGACCAATCTACTCCCTTGCTGCTGTGGACACTGTGGGCGTCCAGGC	60
Db	1	ATGGCTTCCTGACCGCCAGCTCTTGTGCCCTGTGCTGCTTGGCACTTGTGGCACTATCTGCC	60
QY	61	CAGAGCTCGGGGCCCGGGCGGGGCGGGTGTGGCCGCGCGCTATGCGCCCAAGCAAGCTC	120
Db	61	CAGAGCTCGGGGCCCGGGCGAGAACCGGGTGTGGCCGGGGGGTTATGTGGCAAGCAACTT	120
QY	121	GTCGCCGCTACTCTACAGCAATTTGTGTGCCCGGCGCTGCCAGAGCGGACCTTGGCGCCAGT	180
Db	121	GTCGCTCTGCTATACAAACCAATTTGTGTGCCAGTATGCCCGAGCGGACCTTGGGCGCGAGT	180
QY	181	GGCGCAGCGGAGGGGAGGGGTGGCAAGGGGGCTCCAGAGCGTTCCGGGAGCTCGTGGCCCAAC	240
Db	181	GGCGCAGCGGAGGGGAGGGGTAAACAAGGGGGTCTGGAGAGCGCTTCCGGGAGCTCTGATCCCAAC	240
QY	241	TACAAACCCCGCATCATCTTTCAGAGATGAGAGAGAAACAGTGGAGCCGACCGCTGATGACC	300
Db	241	TACAAACCCCGCATATATCTTCAAGSATGAGAGAAACAGCGGCGCAGACCGCTGATGACA	300
QY	301	GAGCGTTTCAGAGAGAGGGGTGAAGGCTTTGGCCATTGCGCGTGAATGAACATGTGGCCCGGA	360
Db	301	GAGCGTTTCAGAAAGAGCGGGGTAAAGCTCTTACCAATCCCGGTGAAGAAATGTGGCCCGGA	360
QY	361	GTCGCGCTACAGATGACTAGAGGGCTGTGGGACAGAGACGGCACACCGCTCAGATTCACCTC	420
Db	361	GTAACCCCTACTGTGTACTAGGAAGCTGGGAGCAGAGAGACGGCACACCGACAGATTCACCTC	420
QY	421	CACATCAGAAAGGCCGTGTGGACATCACTAGTCTGACCGCGACCGCAACAAGTATGGG	480
Db	421	CACATCAGAAAGCCGAGCTTGTGACATCAACACGTCGTCGCGGTGACCGTAAATAGTATGGT	480
QY	481	TTTGCTGGGCGCCCTTCGCAAGTGGAAAGCCGGGCTTCGACTGGGGTCTATCTACAGATCCCGCAAC	540
Db	481	TTGTGGGGCGCCCTTAGCTGTGGAAAGCCGGATTTCACATGGGGTCTATCTACGAGTCCCGCAAC	540
QY	541	CACGCGCAAGTGTGGGTAAAGCTGATTAACCTATGGCGGGTCCGGGGGCGGGGCTGCTTT	600
Db	541	CACATCCACGATGTGGGTAAAGCTGATTAACCTATGGCGGGTCCGAGCGGAGGGTGGCTTT	600
QY	601	CCGGGAATTCGAAGTGTGCGCTGTGGAGCGGCGAGCGGAAGGGGCTGCGGGAATTCAC	660

Db	601	CCGGGAATGCGACAGGTGCGCCCTTCGGAGCGCGGACGGAGAGGGCTTAGGGAACTACAT	650
Qy	661	CGCGAGACTGGGTTTTTGGCGCCGATGCGTCAGGCCGGGTGATGCCACGCCGGTCTG	720
Db	661	CGTGTGACTGGTACGTACGTAGCGCCGCTGATGCAAGCGGCCGAGTGTACCCACGCCACTGCTG	720
Qy	721	CTCTTCCGGAGCCGGGACTTGCAGAGCCCGGGCTCAATTGTGTGGCTGTGAGACCGAGTG	780
Db	721	CTCTTCCTGGAGCCGGGATCTGCAAGCCCGCGCTCTGTGTGTGCTGTGAGACCGAGCG	780
Qy	781	CGTCACGCAAACTGTGTCTACAGCCCTGGACACTGTGTGTCCGCTGAGAGCGCGCG	840
Db	781	CGTCGCGCAAACTGTGTCTCACACCCTGGCACTGTGTGTGCTGTGTGGCGGCGACGG	840
Qy	841	CCCCGCGCAGGCACCTTGTGCACCGGTGTTCGGCGCCGCGCTACGGCTGTGGGACTCGTG	900
Db	841	CGCTCTCCAGGTGACTTGTGCACCGGTGTTCGGCGCCGCGCTTACGTGTGCGACACTCGTG	900
Qy	901	CTGGCGCCCGCGGGGAGTGTGCTTGGAGCAGCGCGGTGGCCCGTGTGGCGGGAGAA	960
Db	901	CTGGCTCCCGCGGGGAGACGGGCTTCAGCCGGCGCGGTGACCCGGGTGGCGGCGAAGA	960
Qy	961	GCCGTGGGCGTGTGCGCGCGCTCACCGCGCACGGAGCGTCTGTGTGAAGATGCTGTG	1020
Db	961	GCCGTGGGCGGTTCGCAACCGCTCACTGGCGACGGAGACCTCTGTGTCAACAGACTCTC	1020
Qy	1021	GCGCTGTGTACGGGGTTCGTGGAGAGTCAACGATGGGGGACACGGGCTTTCGCCCCCTG	1080
Db	1021	GCGCTGTGTACGGGGTTCGTAGAGATCAACGATGGGGGACACCGCGCTTTCGCCCCCTTG	1080
Qy	1081	AGACTGTGTACAGCGGTACGAGGGGCGGTGTCCCGCGGGGCGGTGCACAGCGCATGGCAT	1140
Db	1081	CGGCGTGTACAGCGGTGTGAGGGGCGGTGTCTCCCTGGGGGGTGTCAAGTCAAGCGCATGGCAT	1140
Qy	1141	CATTGTGTACTGTGCGTCTCTTACCGCTTACGGGAGAGACTGTGGGCTG	1190
Db	1141	CATTGTGTACTGTGCGCTTCTTACCGTGTGGCGGAGAGACTGTATGGGCTG	1190

RESULT 9

ID V62397 standard; cDNA; 602 BP.

DT 02-FEB-1999 (first entry)

KW Desert hedgehog; HUHh; human; ds.

FH Key

FT /tag- a

ET / *tag= b

PD 28-OCT-1998.

PR 14-APR-1998;

PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU

DR WPI; 98-544642/47.

PT Human Desert hedgehog protein - and corresponding DNA and monoclonal

PS Example 1-2; Page 27-28; 39pp; English.

CC a precursor form (see W79595) of novel human Desert hedgehog protein

CC originally identified in human following ARH-77 (ATCC CRL-1621) cell

CC V62405-06) yielded cDNA encoding the N-terminal region. DNA

CC W79593-95) of human Desert hedgehog are claimed, as are a monoclonal

CC antibody (Mab) that recognises the protein, a process for producing
CC the protein, and a method for detecting the protein using the Mab.
CC The hedgehog protein, DNA and Mab can be used to elucidate
CC hereditary morphological abnormalities in humans to establish their
CC treatments and diagnoses.

50	Sequence	602 BP,	111 A;	187 C;	199 G;	105 T;
----	----------	---------	--------	--------	--------	--------

Query Match	49.28;	Score 586.4;	DB 1;	Length 602;
Best Local Similarity	99.08;	Pred. No. 1.7e-113;		
Matches 590;	Conservative	0;	Mismatches 6;	Indels 0;

OY	1	ATGGCTATCCGACCAATCTACTGCGCCCTTGAGCTTGAGCACTTGAGCGCTCCAGGC	60
Db	7	ATGGCTCTCTGACCAATCTACTGCGCCCTGTGTGCTTTGGCACTTTTGGCGTCCAGGC	66
OY	61	CAGAGCTCGGGGCGGGGCGGGGGCGGGTGTGGCCGGGCGCGCTATGGCGGCAACAGCTC	120
Db	67	CAGAGCTCGGGGCGGGGCGGGGGCGGGTGTGGCCGGGCGCGCTATGGCGGCAACAGCTC	126
OY	121	GTGGCGGCTACTCTACACCAATTTTGTGGCCGGGCGTGGCAGAGCGGACCTTGGGCGCACT	180
Db	127	GTGGCGGCTACTCTACACCAATTTTGTGGCCGGGCGTGGCAGAGCGGACCTTGGGCGCACT	186
OY	181	GGGCGAGCGAGAGGGGAGGGTGGCAAGGGGGCTCCGAGCGCTTCCGGGACCTGTGCCCAAC	240
Db	187	GGGCGAGCGAGAGGGGAGGGTGGCAAGGGGGCTCCGAGCGCTTCCGGGACCTGTGCCCAAC	246
OY	241	TACAACCCCGCATCATCTTCAAGGATAGGAGAACAAGTGGAGCGCACCGGCTGATAGCC	300
Db	247	TACAACCCCGCATCATCTTCAAGGATAGGAGAACAAGTGGAGCGCACCGGCTGATAGCC	306
OY	301	GAGCGTTGCAAGGAGAGGGGTGAAGCTTTGGGCATTTGCCGTGATGAACATGTGGCCGGA	360
Db	307	GAACTTTTAAGGAACGGGGTGAACGCTTTGGCATTGGCGGTGAAGAACATGTGGCCGGA	366
OY	361	GGGCGCTACGAGTACATGAGGGGCTGGGAGCAAGAGCGGCCACACAGCGTCAGAGTTACTC	420
Db	367	GTGGCGCTACGAGTACATGAGGGGCTGGGAGCAAGAGCGGCCACACAGCGTCAGAGTTACTC	426
OY	421	CACCTACGAAGGCGCGCTTTGGACATCACTACGTCGACCCGAGCCGCAACAATATAGG	480
Db	427	CACCTACGAAGGCGCGCTTTGGACATCACTACGTCGACCCGAGCCGCAACAATATAGG	486
OY	481	TTGCTGGCGCGCTTGCAGTGGAGCCGGGCTTGCAGTGGTCTACTACAGAGTCCCGCAAC	540
Db	487	TTGCTGGCGCGCTTGCAGTGGAGCCGGGCTTGCAGTGGTCTACTACAGAGTCCCGCAAC	546
OY	541	CACGTCACAGTGTGGCTCAAGAGCTATTAACACTAGGCGGTCCGGGCGGGGGGCTG	596
Db	547	CACATTCACAGTGTGGCTCAAGAGCTATTAACACTAGGCGGTCCGGGCGGGGGGCTG	602

RESULT	10
AC	V62398
ID	V62398 standard; cDNA, 575 bp.
AC	V62398:
DT	02-FEB-1999 (first entry)
DE	Human Desett hedgehog protein C-terminal region cDNA
DE	Desett hedgehog; HUH9H; human; ds.
OS	Homo sapiens.
FH	Key
FT	mat_peptide
FT	Location/Qualifiers
FT	2..574
FT	/*tag= a
PN	EP-874048-A2.
PD	28-OCT-1998.
PF	24-APR-1998: 303187.
PR	14-APR-1998: JP-117873.
PR	25-APR-1997: JP-121578.
PA	(HAYB) HAYASHIBARA SHIBITSU KAGAKU.
PI	AtiYasu T, Nakamura S, Ofita K;
DR	WPI: 98-544642/47.
DR	P-PSDB: W79598.

PT Human α -desert hedgehog protein and corresponding DNA and monoclonal
 Pr antibody
 PS Example 1-3; Page 28-29; 39pp; English.
 CC This cDNA sequence codes for the C-terminal region (see W79598) of
 CC a precursor form (see W79594) of novel human desert hedgehog protein
 CC (see also W79593). Human desert hedgehog protein (HdHh) was
 CC originally identified in human following AHH-77 (ATCC CRL-1621) cell
 CC line following RT-PCR amplification screenings. Further PCR (see
 CC W62407-08) yielded cDNA encoding the C-terminal region. DNA
 CC molecules (see W63393-95) encoding mature and precursor forms (see
 CC W79593-95) of human desert hedgehog are claimed, as are a monoclonal
 CC antibody (Mab) that recognises the protein, a process for producing
 CC the protein, and a method for detecting the protein using the Mab.
 CC The hedgehog protein, DNA and Mab can be used to elucidate
 CC hereditary morphological abnormalities in humans to establish their
 CC treatments and diagnoses.
 SQ Sequence 575 bp; 66 A; 183 C; 221 G; 105 T;

Query Match	48.1%;	Score 573.4;	DB 1;	Length 575;
Best Local Similarity	99.8%;	Pred. No. 8.2e-111;		
Matches 574;	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0;

OY	549	CGTTCGGTCAAGACTGATTAATCACTATGGGGGATCCGGGCGGGCGCTCTTCCGGGAAA	608
Db	1	CGTTCGGTCAAGACTGATTAATCACTATGGGGGATCCGGGCGGGCGCTCTTCCGGGAAA	60
OY	609	TGCAACTGTGTGGCTGTGAGCGCGGACGCGGAAAGGGCTGTGGGGAGATTCACCCGCGAGA	668
Db	61	TGCAACTGTGTGGCTGTGAGCGCGGACGCGGAAAGGGCTGTGGGGAGATTCACCCGCGAGA	120
OY	669	CTGGGTTTTGGCGGGCGGATCGTCAAGCGCGGGTGGTGTGCCACGGCGGTCTGTCTTCT	728
Db	121	CTGGGTTTTGGCGGGCGGATCGTCAAGCGCGGGTGGTGTGCCACGGCGGTCTGTCTTCT	180
OY	729	GGACCGGGACTTGGAGCGCGGGGCTTCATTGTGGCTGTGAGACCGAGTGGACTCCACG	788
Db	181	GGACCGGGACTTGGAGCGCGGGGCTTCATTGTGGCTGTGAGACCGAGTGGACTCCACG	240
OY	789	CAAACTGTGTCTACGCGCCCTGGGACCTGTGTGTTCGCGCTCGAGGGCGCGCGCCGCGCC	848
Db	241	CAAACTGTGTCTACGCGCCCTGGGACCTGTGTGTTCGCGCTCGAGGGCGCGCGCCGCGCC	300
OY	849	AGGGGACTTTGACACCGGTGTTCGCGCGCGCGGCTACCGCGCTGGGGACTGGGTCTTGCGCC	908
Db	301	AGGGGACTTTGACACCGGTGTTCGCGCGCGCGGCTACCGCGCTGGGGACTGGGTCTTGCGCC	360
OY	909	CGCGGGGATCCGCTTGGGCGACGCGCGGGCGCGCTGTGGGCGCGGAGGAAGCGGTGG	968
Db	361	CGCGGGGATCCGCTTGGGCGACGCGCGGGCGCGCTGTGGGCGCGGAGGAAGCGGTGG	420
OY	969	CGTGTTCGCGCGCTTCACCGCGACGAGGAGCGTGTCTGTGAACGATGTCTTGCCCTCTTG	1028
Db	421	CGTGTTCGCGCGCTTCACCGCGACGAGGAGCGTGTCTGTGAACGATGTCTTGCCCTCTTG	480
OY	1029	CTACGCGGTTCTGGAGAGTACCAAGTGGGCGCACCGGCTTTTGCCCCCTTGAGACTCT	1088
Db	481	CTACGCGGTTCTGGAGAGTACCAAGTGGGCGCACCGGCTTTTGCCCCCTTGAGACTCT	540
OY	1089	GCACGCGCTAGGGCGCTGTCTCCCGCGGGCGCGG 1123	
Db	541	GCACGCGCTAGGGCGCTGTCTCCCGCGGGCGCGG 575	

RESULT	11
1D	V62396 standard; cDNA; 548 BP.
V62396	
AC	V62396; standard; cDNA; 548 BP.
DT	02-FEB-1999 (first entry)
DE	Human Desert hedgehog protein cDNA clone pHDH/#20
KW	Desert hedgehog; HbDHH; human; ds.
OS	Homo sapiens.
FN	Key Location/Qualifiers

QY 367 CTACGAGTACTGAGGCGTGGGAGCAGAGCGCCACCACGCTAGGATTCCTCCACTAC 426
| | | | |
DB 301 CTACGAGTACTGAGGCGTGGGAGCAGAGCGCCACCACGCTAGGATTCCTCCACTAC 360
| | | | |
QY 427 GAAGGCGGCTCTTTGGACATCACTAGCTGACCGGCGGCAACAAGTGTGGTCTG 486
| | | | |
DB 361 GAAGGCGGCTCTTTGGACATCACTAGCTGACCGGCGGCAACAAGTGTGGTCTG 420
| | | | |
QY 487 GCGCGCTGCGTCAAGGCGGCTTGCAGTGGGTTCTACTACGAGTCCGCAACCAAGTC 546
| | | | |
DB 421 GCGCGCTGCGTCAAGGCGGCTTGCAGTGGGTTCTACTACGAGTCCGCAACCAAGTC 480
| | | | |
QY 547 CACGTGCTGCTCAAGGCTGATTAATCACTAGCTGCGGCGGCGGCGGCGGCG 594
| | | | |
DB 481 CACGTGCTGCTCAAGGCTGATTAATCACTAGCTGCGGCGGCGGCGGCGGCGGCG 528
| | | | |
RESULT 13
ID X16188 standard; cDNA; 1622 BP.
AC X16188:
DE 29-APR-1999 (first entry)
KW Human inh hedgehog cDNA sequence.
KW Patched; hedgehog; ptc therapeutic; neuroprotective; neuronal cell;
KW brain infarction; cerebral infarction; transient ischaemic attack;
KW stroke; cerebral infarct volume; spinal cord; oedema; trauma;
KW haemorrhage; encephalomyelitis; coronary bypass; cerebral hypoxia; ds.
OS Homo sapiens.
FH key Location/Qualifiers
FT CDS 51..1286
FT /tag- a
FT /product- "hedgehog sequence"
PD MO9900117-2.
PD 07-JUN-1999.
PE 26-JUN-1998; 013387.
PR 27-JUN-1997; US-883656.
PA (ONTO-) ONTOGENY INC.
PI Mahantappa NK.
DR WPI: 99-095458/08.
DR P-PSDB; M94474.
PT Method for limiting damage to neurons caused by ischaemic or epoxic
PT conditions - is used for the treatment and prevention of e.g.
PT cerebral infarction, stroke and transient ischaemic attacks
PS Disclosure: Page 74-76: 104pp: English.
CC A method has been developed for limiting the damage to neuronal cells by
CC ischaemic or epoxic conditions by administering a ptc (patched)
CC therapeutic agent to reduce cerebral infarct volume (CIV). Damage to
CC neuronal cells can also be limited by administering a gene activation
CC construct which recombinates with the genomic hedgehog gene to provide a
CC heterologous transcription regulator linked to the coding region of this
CC gene. Administration of the ptc therapeutic agent is used to protect
CC cerebral tissues against ischaemic injury; to treat cerebral infarct or
CC ischaemia, stroke (thrombotic or embolic) and transient ischaemic
CC attacks. It may also be used as a prophylactic in many other cases of
CC injury to the brain or spinal cord, oedema caused by trauma, haemorrhage
CC and encephalomyelitis, or in conjunction with (coronary bypass) surgery.
CC Treatment (which may be prophylactic) is used where ischaemic/epoxic
CC conditions may cause cerebral hypoxia, or progressive loss of neurons
CC due to oxygen depletion, including in patients with hypotension. The
CC treatment reduces CIV by at least 25, particularly at least 70%. The
CC present sequence encodes a hedgehog sequence given in the present
CC invention.
SQ Sequence 1622 BP: 277 A; 549 C; 510 G; 286 T;

QY 98 GCGCTATGCGCGCAAGACAGCTGCTGCGCTACTCTACAAGCAATTTGCCCCGGGTGC 157
| | | | |
DB 163 GCGCAGCGGCAAGCA--AACTGCTGCGGTGCTCTACAAGCAATTTGCCCCGAATGTGC 219
| | | | |
QY 158 CAGAGGAGACCTTGGGCGGCGGAGTGGGCGCAGCGGAGGAGGTGGCAGAGGCTCCGAGC 217
| | | | |
DB 220 CCGAGAGAGACCTTGGGCGGCGGAGGAGGAGGTGGCAGAGGAGTGGCAGGCTCCGAGC 279
| | | | |
QY 218 GCTTCCGGGAGCTGCTGCGGCACTACAACCGGCACTCAATCTTCAAGGATGAGAGACA 277
| | | | |
DB 280 GCTTCAAGAGCTGCTGCGGCACTACAACCGGCACTCAATCTTCAAGGAGGAGAGACA 339
| | | | |
QY 278 GTGAGCGGCGGCGGCGGCTGATGAGCGGAGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
DB 340 CAGGCGGCGGCGGCGGCTGATGAGCGGAGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
QY 338 CCGTATGAGACATGAGGCGGCGGAGTGGGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
DB 400 CCGTATGAGACATGAGGCGGCGGAGTGGGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
QY 398 GCGCAGCGGCGGCGGCTGATGAGCGGAGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
DB 460 GCGCAGCGGCGGCGGCTGATGAGCGGAGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
QY 458 ACCGCGGCGGCGGCGGCTGATGAGCGGAGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
DB 520 ACCGCGGCGGCGGCGGCTGATGAGCGGAGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
QY 518 GCGTATGAGACATGAGGCGGCGGAGTGGGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
DB 580 GCGTATGAGACATGAGGCGGCGGAGTGGGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
QY 578 GCGTATGAGACATGAGGCGGCGGAGTGGGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
DB 640 CAGCAGGCGGCGGCGGCTGATGAGCGGAGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
QY 638 GGAAGGCGGCGGCGGCTGATGAGCGGAGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
DB 700 GCGTATGAGACATGAGGCGGCGGAGTGGGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
QY 698 GCGTATGAGACATGAGGCGGCGGAGTGGGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
DB 760 GCGCAGCGGCGGCGGCTGATGAGCGGAGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
QY 758 TGTGCTGAGGAGCGGCGGAGTGGGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
DB 820 TCCAGGCGGCGGCGGCTGATGAGCGGAGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
QY 818 TGTGCTGAGGAGCGGCGGAGTGGGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
DB 880 TGTGCTGAGGAGCGGCGGAGTGGGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
QY 878 GCGTATGAGACATGAGGCGGCGGAGTGGGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
DB 940 ACCTGAGCGGCGGCGGCTGATGAGCGGAGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
QY 938 TGGCGGCGGCGGCGGCTGATGAGCGGAGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
DB 1000 TGGCGGCGGCGGCGGCTGATGAGCGGAGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
QY 998 GCGTATGAGACATGAGGCGGCGGAGTGGGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
DB 1060 CAGTATGAGACATGAGGCGGCGGAGTGGGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
QY 1058 GCGCAGCGGCGGCGGCTGATGAGCGGAGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
DB 1120 CTAGGAGCGGCGGCGGCTGATGAGCGGAGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
QY 1118 GCGCAGCGGCGGCGGCTGATGAGCGGAGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |
DB 1171 GCTGAGCGGCGGCGGCTGATGAGCGGAGCGGCTTCAAGGAGGAGGTGAACGCTTTGGCATTG 337
| | | | |

QY 1178 AGCTACTGG 1186
Db 1231 TCCTGCTAG 1239

RESULT 14
X07269
ID X07269 standard; cDNA; 1622 BP.
AC X07269;
DT 21-MAY-1999 (first entry)
DE Human Indian hedgehog (Ihh) cDNA.
KW Indian hedgehog; Ihh gene; human; dopaminergic; GABA-ergic;
KW Ptc (patched); patched; signal transduction; Parkinson's disease;
KW Huntington's disease; amyotrophic lateral sclerosis;
KW cerebral ischemia; hypoxia; neuroprotective; gene therapy; ss; ds.
OS Homo sapiens.
FH Key location/qualifiers
FT CDS 51..1286 /*tag= a

MO9904775-A2.
PD 04-FEB-1999. U15419.
PR 24-JUL-1998: US-900220.
PA (ONTO-) ONTOGENE INC.
PI Mananthappa NK, Miao N, Pang K, Wang M;
PI WPI; 99-142578/12.
DR P-PSDB; W97763.
PT Increasing the survival of neuronal, dopaminergic and GABA-ergic
PT cells - by using a ptc therapeutic such as a protein kinase
PT inhibitor or an agent derived from hedgehog polypeptides, useful in
PT the treatment of Parkinson's disease
PS Disclosure: Page 95-97; 138pp; English.
CC This nucleotide sequence comprises a coding region for the human
CC Ihh Indian hedgehog protein (see W97763). The invention is based on
CC the finding that hedgehog proteins are useful as protective agents
CC in the treatment and prophylaxis of neurodegenerative disorders
CC resulting from the loss of dopaminergic and/or GABA-ergic neurons,
CC or the general loss of tissue from the substantia nigra.
CC Exemplary disorders include Parkinson's disease, Huntington's
CC disease (both claimed), amyotrophic lateral sclerosis and cerebral
CC ischemia. The invention relates to hedgehog therapeutics (i.e.
CC hedgehog polypeptides and gene therapy constructs e.g. constructs
CC encoding recombinant hedgehog polypeptides and trans-activation
CC constructs for altering hedgehog gene regulatory sequences) and
CC ptc therapeutics (i.e. agents which mimic the effect of naturally
CC occurring hedgehog proteins on patched signaling) that are
CC effective in both human and animal subjects. A bioactive
CC polypeptide comprising amino acid residues 28-202 of human Ihh is
CC preferred. The products can also be used for the maintenance of
CC differentiated neurons in cultures, and to enhance the implantation
CC of such neuronal cells in an animal. They can also be used to
CC prevent or treat neurodegenerative conditions arising from the use
CC of certain drugs, and in the prevention and/or treatment of hypoxia,
CC e.g. as a neuroprotective agent.
CC Sequence 1622 BP; 277 A; 549 C; 510 G; 286 T;

Query Match 38.1%; Score 453.8; DB 1; Length 1622;
Best Local Similarity 63.5%; Pred. No. 5.4e-86;
Matches 730; Conservative 0; Mismatches 407; Indels 12; Gaps 2;

QY 38 TGGCACTTCTGGCGCTGCCAGCCAGAGCTGGCGGCGCGGCGGCGGTGGCCGCG 97
Db 103 TGCTCTGCTGGTGGTCCCGCGCATGGGGTGGCGGCGGTGGTGGGCGACCGCC 162

QY 98 GCCGTATGCGCGGCAAGAGCTGCGCGCTACTACAGCAATTTGTGCCGCGTGC 157
Db 163 GCGCGACCGCCAGCA--AACTCGCGCGCTGCGCTACAGAGCAAGTCCACCAATGTGC 219

QY 158 CAGACGAGACCTGGGCGCAGTGGCCAGCGAGGGAGGTGGGAGGGGCTCCGAGC 217
Db 220 CCGAGAAACCTTGGGCGCGCGAGCGATGAGGCAAGATCCTCGCAGCTCCGAGC 279

QY 218 GCTTCGGGACCTGTGCCCACTACAAACCCGACATCATCTTCAAGATGAGAGACA 277
Db 280 GCTTCAGAGACTCACCCCAATTACATCCAGACATCATCTTCAAGAGAGAGACA 339

QY 278 GTGAGCCGACCCGCTGATGACCGAGCTTGCAAGAGAGAGGTAAAGCTTTGGCATTG 337
Db 340 CAGCGCGCGACCGCTTATGACCCGAGCTGCAAGAGACCGCTCAACTCGTGTCTACT 399

QY 338 CCGTATGATGACATGTGGCGCGGAGTGGCGCTACGAGTGAAGAGGCTGGAGCAGAGACG 397
Db 400 CGGTATGATGACATGTGGCGCGGAGTGGAGAGTGGCGGCTGACGAGAGGCTGGAGCAGAGACG 459

QY 398 GCGACACGCTCAGATTCATCTCCACTACAGAGCGCGCTTTGGAGATCACTACGTCTG 457
Db 460 GCGACACCTCAGAGAGAGTCCCTGATATGAGCGCGCGGCTGAGATCACCACATCAG 519

QY 458 ACCCGACCGCAAAATATATGAGTGTGCTGGCGCGCTCGAGTGAAGCCGCGCTCACT 517
Db 520 ACCCGACCGCAAAATATATGAGTGTGCTGGCGCGCTCGAGTGAAGCCGCGCTTCACT 579

QY 518 GGGTCTACTACGAGTCCCGCAACGAGCTCCAGTGTGCTCAAGTGTATCACTACTGG 577
Db 580 GGGTCTACTACGAGTCCAAAGGCGCGAGCTGATTCCTGCTCAAGTGTGAGAGCTTGGCGCG 639

QY 578 CGGTCCGGCGCGGCGCTCTTCCGGGAAATGCAACTGTGCGCTGTGAGCGCGGAGC 637
Db 640 CAGCGACGAGCGGCGGCTCTTCCCTCGGAGCGCGAGTACGCTGAGAGTGTGGGCGCG 699

QY 638 GGAAGGCGCTCGGAGACTGACCGCGGAGACTGTTTGGCGCGCGAGTGCCTACGCC 697
Db 700 GTGTGGCTTGTACCGCTGAGGCGGAGACCGTGTGCTGCTGATGGGAGAGATGGA 759

QY 698 GGGTGTGCGCCACCGCGCTGCTCTTCCGACCGGAGCTTCCAGCGCGGCGCTCAT 757
Db 760 GCGCGACCTTACGAGTGTGCTTCTTCTGAGCCGCGAGCCCAAGAGCTGAGAGCT 819

QY 758 TTGTGCTGTGAGACGAGTGGCTTCCAGCAAACTGTGCTGACCGCTGACCTGCG 817
Db 820 TCCAGGTCATGAGACATCAGAGACCGCGCGCTGAGCTACACACCGCTACCTGC 879

QY 818 TGTGTGCGCTCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGCGCGCG 877
Db 880 TCTTTAGCGTGAACATCAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 939

QY 878 GGTACCGCGTGGGAGTGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 937
Db 940 ACGTGACCGCTGGCGAGTACGTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 999

QY 938 TGGCGCGTGGCGCGGAGAGACCGTGGCGGTGTCGCGCGCGCTACCGCGCAGCGGA 997
Db 1000 TGGCAGCTGTCTACACAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1059

QY 998 CGCTGCTGTGTAAGATGTCTGCGCTTGTGCTGAGGCGGTGTAAGATGACAGTGGG 1057
Db 1060 CACTGTGTGTGAGATGTGTGTGATCTGCTTCCGCGCGCGCGCGCGCGCGCGCGCG 1119

QY 1058 CGCACCGCGCTTGTGCCCGCTTGAGACTGCTGACGCGCGCTAGGGCGCGTGTCCCGGCG 1117
Db 1120 CTGATGTGCGCTTGTGGCGCGCTTGAGACTTTCACAGCTTGGCATGG-----GGCA 1170

QY 1118 GGGCGGTCCAGCCGACATGTCATGATGATGATGATGATGATGATGATGATGATGATG 1177
Db 1171 GCTGAGACCGCGGAGAGGTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1230

QY 1178 AGCTACTGG 1186
Db 1231 TCCTGCTAG 1239

RESULT 15
X25104
ID X25104 standard; cDNA; 1622 BP.
AC X25104;

DT 05-JUL-1999 (first entry)
 DE Human Indian hedgehog protein Ihh cDNA.
 KW Indian hedgehog; Ihh gene; human; hedgehog therapeutic;
 KW ptc therapeutic; patched; signal transduction; muscle atrophy;
 KW cachexia; muscular myopathy; myoblastic sarcoma; therapy; ds.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 51..1286
 FT /tag a
 PN MO9910004-A2.
 PD 04-MAR-1999.
 PF 28-AUG-1998; 017922.
 PR 29-AUG-1997; 05-057394.
 PA (ONTO-) ONTOGENY INC.
 PI Bladgen CS, Currie PD, Hughes SM, Ingham PW;
 DR WPI: 99-243557/20.
 P-PSDB: X05516.
 PT A new method to regulate muscle growth.
 PS Disclosure: Page 104-106; 130pp; English.
 CC This nucleotide sequence comprises a coding region for the human
 CC Indian hedgehog protein Ihh (see X05516). The invention relates to
 CC a method for modulating the formation and/or maintenance of muscle
 CC tissue by ectopically contacting muscle cells, especially muscle
 CC stem/progenitor cells, in vitro or in vivo, with a hedgehog
 CC therapeutic (i.e. hedgehog polypeptides and gene therapy
 CC constructs) or ptc therapeutic (i.e. a small organic molecule that
 CC mimics the effect of hedgehog proteins on patched signaling, or
 CC activates or potentiates patched signaling) in an amount effective
 CC to alter the growth state of the treated cells. Also claimed is a
 CC method for treatment or prevention of disorders of, or surgical or
 CC cosmetic repair of, such muscle tissues, by administering a
 CC hedgehog polypeptide or ptc therapeutic. The disorder may be
 CC muscle atrophy, in particular skeletal muscle atrophy or cardiac
 CC muscle atrophy, cachexia, or muscular myopathy (all claimed). The
 CC hedgehog polypeptide or ptc therapeutic can inhibit growth of
 CC myoblastic-derived tissue to provide treatment of hyperplastic or
 CC neoplastic growth of muscle tissue such as in myoblastic sarcoma
 CC (also claimed). The hedgehog therapeutic preferably comprises at
 CC least a bioactive extracellular portion of a hedgehog protein (see
 CC X05510-19) encoded by a vertebrate hedgehog gene (see X25098-107),
 CC especially a human hedgehog gene.
 SO Sequence 1622 bp; 277 A; 549 C; 510 G; 286 T;

Query Match 38.1%; Score 453.8; DB 1; Length 1622;
 Best Local Similarity 63.5%; Pred. No. 5.4e-86;
 Matches 730; Conservative 0; Mismatches 407; Indels 12; Gaps 2;

QY 38 TGGCACTTCTGGCGCTGCGCCAGAGAGTGGGGGCGGGGGCGCGTGGCGGC 97
 DB 103 TGT 162

QY 98 GCGGCTATGGCGCGCAAGACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 157
 DB 163 GCGGCGCGCGCGCAAGACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 219

QY 158 CAGAGCGGAGCGCTGGGGCGAGTGGCGCGAGCGGAGGAGGAGGAGGAGGAGGAGG 217
 DB 220 CCGAGAGAGAGCGCTGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 279

QY 218 GCTTCCGGGAGCTGT 277
 DB 280 GCTTCAAGAGAGT 339

QY 278 GTGGAGCGGAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 337
 DB 340 CAGGCGCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 399

QY 338 CCGTGTATGAGCATGTGGCGCGAGTGGCGCTTACGAGAGTGTGTGTGTGTGTGTGT 397
 DB 400 CGGTGTATGAGCATGTGGCGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 459

QY 398 GCCACACGCTCAGATTCACTCACTAGAGAGGCGGTGTGTGTGTGTGTGTGTGTGT 457

DB 460 GCCACACTCAGAGAGTCCCTTCATTATGAGGCGCGCGGGGAGACATCACCATCAG 519
 QY 458 ACCGCGACCGCAACAGTATGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 517
 DB 520 ACCGCGACCGCAACAGTATGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 579

QY 518 GGGTCTACTAGAGT 577
 DB 580 GGGTGTATATAGAGT 639

QY 578 CGGTCCGGGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 637
 DB 640 CAGCCAAAGAGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 699

QY 638 GGAAGGCGTGGCGGGAACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 697
 DB 700 GTGTGCTGT 759

QY 698 GGGT 757
 DB 760 GCCCACTTACAGAGT 819

QY 758 TTGT 817
 DB 820 TCCAGT 879

QY 818 TGT 877
 DB 880 TGT 939

QY 878 GGT 937
 DB 940 AGCTGT 999

QY 938 TGGCCCGT 997
 DB 1000 TGGCAGT 1059

QY 998 CGGT 1057
 DB 1060 CACTGT 1119

QY 1058 CGCACGCGCTTGT 1117
 DB 1120 CTAGT 1170

QY 1118 GGGCCGT 1177
 DB 1171 GCTGT 1230

QY 1178 AGCTGT 1186
 DB 1231 TCCTGT 1239

Search completed: June 5, 2000, 07:58:49
 Job time: 3814 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2000, 08:17:52 ; Search time 22.84 Seconds
(without alignments)
1016.523 Million cell updates/sec

Title: US-08-900-220-17

Sequence: 1 MALNTNLPCCCLALLALPA.....PTGMHYSRLYRLAEELG 396

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 168808 segs, 58629743 residues

Total number of hits satisfying chosen parameters: 168808

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : PIR_63:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2008	96.2	396	B49425	Desert hedgehog pr
2	1217	58.3	437	A49425	Sonic hedgehog pro
3	1203	57.6	437	B53193	hedgehog homolog v
4	1198	57.4	425	A49424	patterning protein
5	1120	53.6	418	A53193	hedgehog homolog v
6	1092.5	52.3	444	S56765	morphogen Xhh prec
7	1069	51.2	415	A49426	sonic hedgehog gen
8	956.5	45.8	336	C49425	Indian hedgehog pr
9	951	45.5	471	A46400	segment polarity p
10	495	23.7	94	G02735	desert hedgehog -
11	209	10.0	868	T22281	hypothetical prote
12	194.5	9.3	1226	T24045	hypothetical prote
13	193.5	9.3	1021	T23252	hypothetical prote
14	186.5	8.9	1207	T23754	hypothetical prote
15	182	8.7	481	T27665	hypothetical prote
16	163	7.8	484	T34504	hypothetical prote
17	161.5	7.7	629	T19563	hypothetical prote
18	150.5	7.2	615	T29550	hypothetical prote
19	135	6.5	205	T26220	hypothetical prote
20	117	5.6	3172	S22012	erythronolide synt
21	115.5	5.5	3670	T36249	cya peptide synthe
22	114.5	5.5	687	T08528	probable DNA topoi
23	112.5	5.4	846	Q08EC3	HQRF1 protein - hu
24	111.5	5.3	788	Q08EC3	HQRF1 protein - hu
25	111.5	5.3	3178	S13595	6-deoxyerythronoli
26	106.5	5.1	481	T27975	hypothetical prote
27	103.5	5.0	3391	U50219	polypotein - deng
28	103	4.9	672	T36083	hypothetical prote
29	103	4.9	3519	S43048	polyketide synthas
30	102.5	4.9	3388	GNWVDP	genome polypotein

31	100.5	4.8	3391	1	GNWV26	genome polypotein
32	98	4.7	474	2	C75625	hypothetical prote
33	97	4.6	3391	1	GNWV16	genome polypotein
34	96.5	4.6	541	2	T35143	probable monooxyge
35	96	4.6	759	2	S32875	hypf protein - Rhl
36	96	4.6	3396	1	A42551	genome polypotein
37	95.5	4.6	488	2	D70614	hypothetical prote
38	95.5	4.6	1039	2	A34269	integrin alpha-2b
39	95.5	4.6	1958	2	B40505	hypothetical prote
40	94.5	4.5	376	2	F70576	hypothetical prote
41	94.5	4.5	1763	3	T17465	rifamycin polyketi
42	94	4.5	401	2	B75562	molybdopterin bios
43	94	4.5	811	2	S08579	hypothetical prote
44	93	4.5	1733	1	B45344	probable nuclear a
45	92	4.4	987	2	A75496	conserved hypothet

ALIGNMENTS

RESULT 1
B49425
Desert hedgehog protein precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Dec-1999
C:Accession: B49425
R:Echeillard, Y.; Epstein, D.J.; St-Jacques, B.; Shen, L.; Mohler, J.; McMahon, T.A.; M
Cell 75, 1417-1430, 1993
A:Title: Sonic hedgehog, a member of a family of putative signaling molecules, is imp
A:Reference number: A49425; M01D:94094334
A:Accession: B49425
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-396 <ECHO>
A:Cross-references: GB:X76292; NID:g443941; PIDN:CA53924.1; PID:g443942
C:Genetics:
A:Gene: Dhh
C:Superfamily: sonic hedgehog protein

Query Match	Score	DB 2:	Length
Best Local Similarity	96.2%	Pred. No. 7.4e-153;	396;
Matches 382; Conservative	6;	Mismatches 8;	Indels 0; Gaps 0;
QY	1	MALNTNLPCCCLALLALPAOSCGPGRGVRRYARKOLVPLLYKQFVPCVPERTIGAS	60
DB	1	MALPASLLPCCCLALLALSAOSCGPGRGVRRYARKOLVPLLYKQFVPCVPERTIGAS	60
QY	61	GPAGRVARGSERPRDLVPNYNPDIIFKDEENSGADRLMTERCKERYNALAIAYNMMPG	120
DB	61	GPAGRVARGSERPRDLVPNYNPDIIFKDEENSGADRLMTERCKERYNALAIAYNMMPG	120
QY	121	VRLLVTGMDDEGHHADDSLHYEGRALDITTSDDRRKYGGLAALANAEAGDWMYYESRN	180
DB	121	VRLLVTGMDDEGHHADDSLHYEGRALDITTSDDRRKYGGLAALAEAGDWMYYESRN	180
QY	181	HVHSVKRDNLSAVRAGCGFGNATVRLMSGERGRLRELHGDVWLAAASGRVPTPV	240
DB	181	HIHSVKRDNLSAVRAGCGFGNATVRLMSGERGRLRELHGDVWLAAASGRVPTPV	240
QY	241	LFELRDLDQRRASFVAVETEMPRLTLTPMHLVFAARGPAPADGFAVFAARLRAGDSV	300
DB	241	LFELRDLDQRRASFVAVETEMPRLTLTPMHLVFAARGPAPADGFAVFAARLRAGDSV	300
QY	301	LAPGDALRPARVARVAREEAVGFAPLTAGCTLLVNDVLAACYAVLESQWARRAPAPL	360
DB	301	LAPGDALQPARVARVAREEAVGFAPLTAGCTLLVNDVLAACYAVLESQWARRAPAPL	360
QY	361	RLHRLGALLPGAVOPTGMHYSRLYRLAEELG 396	
DB	361	RLHRLGALLPGAVOPTGMHYSRLYRLAEELG 396	

RESULT 2

A49425

Sonic hedgehog protein precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Dec-1999

C:Accession: A49425

R:Schelard, Y.; Epstein, D.J.; St-Jacques, B.; Shen, L.; Mohler, J.; McMahon, J.A.; McMahon

Cell 75, 1417-1430, 1993

A:Title: Sonic hedgehog, a member of a family of putative signaling molecules, is implied

A:Reference number: A49425; MOTID:94094334

A:Accession: A49425

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-437 <ECH>

A:Cross-references: GB:X76290

A:Note: authors translation is shown for the codon TCC at position 436

C:Genetics:

A:Gene: Shh

C:Superfamily: sonic hedgehog protein

Query Match

58.3%; Score 1217; DB 2; Length 437;
Best Local Similarity 58.7%; Pred. No. 1.1e-89;

Matches 249; Conservative 53; Mismatches 82; Indels 40; Gaps 12;

QY 7 LPLCLALIA-----LPASCGPGRGPRGRRRYARKQVLPILYKQFVGPVPERITLGASG 61

DB 4 LARCFLYIIASSLLVCPGLACGPRG-FGKRHPKK-LTPLAYKQFIPVNAEKTIGASG 61

QY 62 PARGVARGSEPRDLPVNPDIIFKDEENSGADRLMTERCKEKNVNLATAVMMMPGV 121

DB 62 RYEGKITRNSERKELTPNPNPDIIFKDEENTGADRLMTERCKEKNVNLATAVMMMPGV 121

QY 122 RLRYTEGMDGHHADSLHYEGRALDITTSDDRNKYGGLARLAVEAGFDWYVESRNH 181

DB 122 RLRYTEGMDGHHSESLHYEGRAVDITTSDDRSKYGLARLAVEAGFDWYVESKAR 181

QY 182 VHYSVKADNSLAVRAGGCPFGNATVRLMSGERKGLRELHSGDWVLAADASGRVPTPYLL 241

DB 182 IHCSYKAENSVAAKSDGCPFGSATVHLEGGTKLVKDLSPGDVLAADDOGRILYSDFLT 241

QY 242 FLDRDQRRASPVAVETEMPRKLLTPHILVFAA-----RGPAAPGDPAPVFARLRAG 297

DB 242 FLDRDBAKKVFYVETLERERLLTAHLVAPHNDSGFTPPG---SALFASRVPRG 298

QY 298 DSVLA---PGGD-ALRPARAVYA-REAVGVAPLTAHGTLLVNDVLAACYAVLESQW 352

DB 299 QRYVVAERGGDRRLPAVHSTVLRREAGAVAPLTAHGTLLINRYLASCYAVIEHSW 358

QY 353 AHRAFAFRLIHALAALAPARTDGGGGSLPAPASATEARGAEPAGIMYSQLLYHIG 418

DB 359 AHRAFAFRLIHALAALAPARTDGGGGSLPAPASATEARGAEPAGIMYSQLLYHIG 418

QY 392 EELL 395

DB 419 TWLL 422

RESULT 3

B53193

hedgehog homolog vhh-1 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Dec-1999

C:Accession: B53193

R:Roelink, H.; Augsburger, A.; Heemskerk, J.; Korch, V.; Norlin, S.; Ruiz i Altaba, A.;

Cell 76, 761-775, 1994

A:Title: Floor plate and motor neuron induction by vhh-1, a vertebrate homolog of hedgeh

A:Reference number: A53193; MOTID:94170375

A:Accession: B53193

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-437 <ROE>

A:Cross-references: GB:I27340; NID:g452122; PIDN:AAA20999.1; PID:g452123

C:Superfamily: sonic hedgehog protein

Query Match

57.6%; Score 1203; DB 2; Length 437;
Best Local Similarity 57.8%; Pred. No. 1.5e-88;

Matches 245; Conservative 52; Mismatches 87; Indels 40; Gaps 10;

QY 7 LPLCLALIA-----ALALPQSCGPRGPRGRRRYARKQVLPILYKQFVGPVPERITLGASG 61

DB 4 LARCFLYIIASSLLVCPGLACGPRG-FGKRHPKK-LTPLAYKQFIPVNAEKTIGASG 61

QY 62 PARGVARGSEPRDLPVNPDIIFKDEENSGADRLMTERCKEKNVNLATAVMMMPGV 121

DB 62 RYEGKITRNSERKELTPNPNPDIIFKDEENTGADRLMTERCKEKNVNLATAVMMMPGV 121

QY 122 RLRYTEGMDGHHADSLHYEGRALDITTSDDRNKYGGLARLAVEAGFDWYVESRNH 181

DB 122 RLRYTEGMDGHHSESLHYEGRAVDITTSDDRSKYGLARLAVEAGFDWYVESKAR 181

QY 182 VHYSVKADNSLAVRAGGCPFGNATVRLMSGERKGLRELHSGDWVLAADASGRVPTPYLL 241

DB 182 IHCSYKAENSVAAKSDGCPFGSATVHLEGGTKLVKDLSPGDVLAADDOGRILYSDFLT 241

QY 242 FLDRDQRRASPVAVETEMPRKLLTPHILVFAA-----RGPAAPGDPAPVFARLRAG 297

DB 242 FLDRDBAKKVFYVETLERERLLTAHLVAPHNDSGFTPPG---SALFASRVPRG 298

QY 298 DSVLA---PGGD-ALRPARAVYA-REAVGVAPLTAHGTLLVNDVLAACYAVLESQW 352

DB 299 QRYVVAERGGDRRLPAVHSTVLRREAGAVAPLTAHGTLLINRYLASCYAVIEHSW 358

QY 353 AHRAFAFRLIHALAALAPARTDGGGGSLPAPASATEARGAEPAGIMYSQLLYHIG 418

DB 359 AHRAFAFRLIHALAALAPARTDGGGGSLPAPASATEARGAEPAGIMYSQLLYHIG 418

QY 392 EELL 395

DB 419 TWLL 422

RESULT 4

A49424

patterning protein Sonic hedgehog precursor - chicken

C:Species: Gallus gallus (chicken)

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Dec-1999

C:Accession: A49424

R:Riddle, R.D.; Johnson, R.L.; Lauter, E.; Tablin, C.

Cell 75, 1401-1416, 1993

A:Title: Sonic hedgehog mediates the polarizing activity of the zpa.

A:Reference number: A49424; MOTID:94094333

A:Accession: A49424

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-425 <RID>

A:Cross-references: GB:I28099; NID:g453526; PID:g453527

C:Superfamily: sonic hedgehog protein

F:1-26/Domain: signal sequence #status predicted <SIG>

Query Match 57.4%; Score 1198; DB 2; Length 425;

Best Local Similarity 59.4%; Pred. No. 3.6e-88;

Matches 244; Conservative 56; Mismatches 91; Indels 20; Gaps 10;

QY 1 MLTLNLDP---LCCIALALPQSCGPRGPRGRRRYARKQVLPILYKQFVGPVPERITL 57

DB 4 MLTLTILLVGIC--ALVSSGLICGPRG-IGKRHPKK-LTPLAYKQFIPVNAEKTIL 59

QY 58 GASGARGVARGSEPRDLPVNPDIIFKDEENSGADRLMTERCKEKNVNLATAVMM 117

DB 60 GASGARGKITRNSERKELTPNPNPDIIFKDEENTGADRLMTERCKEKNVNLATAVMM 119

QY 118 MGVRLRYTEGMDGHHADSLHYEGRALDITTSDDRNKYGGLARLAVEAGFDWYVE 177

A: Introns: 41/1; 79/3; 131/3; 197/1; 262/1; 297/3; 566/3; 610/3; 735/3; 775/3; 834/3; 88

Query Match 9.3%; Score 193.5; DB 2; Length 1021;

Best Local Similarity 27.1%; Pred. No. 1.6e-07;

Matches 59; Conservative 41; Mismatches 83; Indels 35; Gaps 8;

QY 199 CFPGNATVRLMSGERKGLREHGRDVLADASGRVPPVLLFLDRDQRRASFVAVT 258

DB 802 CFSRDLWTTPSG-KKRMDEIEIGDVLADLTALF-SAITLMHREPEVQEFLEIKT 859

QY 259 EMPPRKLLTPMHLVFAAR-----GPAP-----ACDPAPYFARLRAG 297

DB 860 D-NGKTLQLAGFITATCRRLPSSNSSLNSTPERYHLIDTLPDSETKLASQKIG 918

QY 298 DSVLAPGDALPARVARYAREAVGFAPLTAGTLLVNDVLAACVAVLESHQ----- 352

DB 919 ECLLHNGQFNMOKIDTSKIVSTIGISPLTENGRIIVNDVLAACVSEVQONVLOTTF 978

QY 352 WAHRAFAPIR--LHAALGALLPEGAVOPTGMWYSRL 387

DB 979 W--AFDRRLNLIQYFGDLYDELPYGTSLYKEVL 1013

RESULT 14

123754 hypothetical protein T05C12.10 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999

C:Accession: T23754; T24513

R:Thomas, K.

Submitted to the EMBL Data Library, June 1995

A:Reference number: 219793

A:Accession: T23754

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1207 <WIL>

A:Cross-references: EMBL:249968; PIDN:CAA90265.1; GSPDB:GN00020; CESP:T05C12.10

A:Experimental source: clone M110

R:Burton, J.

Submitted to the EMBL Data Library, October 1995

A:Reference number: 219901

A:Accession: T24513

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1207 <W12>

A:Cross-references: EMBL:266500; PIDN:CAA91313.1; GSPDB:GN00020; CESP:T05C12.10

A:Experimental source: clone T05C12

C:Genetics:

A:Gene: CESP:T05C12.10

A:Map position: 2

A:Introns: 31/3; 87/2; 141/3; 180/2; 203/3; 267/1; 776/2; 794/2; 834/2; 1086/3; 1143/1;

Query Match 8.9%; Score 186.5; DB 2; Length 1207;

Best Local Similarity 31.9%; Pred. No. 7e-07;

Matches 60; Conservative 36; Mismatches 71; Indels 21; Gaps 8;

QY 185 SYKANSLAVRAGG---CFPGNATVRLMSGERKGLREHGRDVLADASGRVPPV 239

DB 942 AVLATPGAGAGGSRSCFASDLVTTVQKR-MDELQIGDYLVPPSSGNVLYKEYK 1000

QY 240 LFLDRDQRRASFVAVENTEMPRKLLTPMHLVFAARGA-----PAPGDA---PV 289

DB 1001 EMFYHREPRETFNFVLYTK-SGRKLSLGRHLLPYAECSEQVEQYTMNPDGIDVAMRESK 1059

QY 290 FARRLRAGDSVLA--PGGDALPARVARYAREAVGFAPLTAGTLLVNDVLAACVAVL 347

DB 1060 VAEKARKGCEVLSIDSEGVAD-EIVRVGMTNVGIYSPMTVEGSLIYDVLSGCSFSL 1118

QY 348 ESHQWHR 355

DB 1119 ESHS-AHK 1125

RESULT 15

127665 hypothetical protein ZK1037.10 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T27665

R:Basham, V.

Submitted to the EMBL Data Library, October 1996

A:Reference number: 220401

A:Accession: T27665

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-481 <WIL>

A:Cross-references: EMBL:281142; PIDN:CAB03509.1; GSPDB:GN00023; CESP:ZK1037.10

A:Experimental source: clone ZK1037

C:Genetics:

A:Gene: CESP:ZK1037.10

A:Map position: 5

A:Introns: 37/3; 75/3; 141/3; 173/2; 253/1; 298/3; 310/1; 371/2; 424/3

Query Match 8.7%; Score 182; DB 2; Length 481;

Best Local Similarity 29.6%; Pred. No. 5.2e-07;

Matches 47; Conservative 31; Mismatches 73; Indels 8; Gaps 4;

QY 199 CFPGNATVRLMSGERKGLREHGRDVLADASGR-VVPTPVLLFLDRDQRRASFVAVT 257

DB 314 CFPNDVAVNVEKAVKRMDELEIGDWBALDENGEDITFLPVKTYLHRDPQEAFFLEFS 373

QY 258 TEMPRKLLTPMHLVFAA---RGAPAPGDAVFARLRAGDSVLA PGDAL---RPA 311

DB 374 LD-NGEFTLREKHLYVTECRQNSSELKISWESISAGKVNAAGCFYLAQSEALTKRYLV 432

QY 312 RVARYAREAVGFAPLTAGTLLVNDVLAACVAVLESH 350

DB 433 ELDIKRVKKTGIYAPMTSGHLLVNNKIHTSCHSEVDH 471

Search completed: June 5, 2000, 08:17:55
Job time: 1235 sec

THIS PAGE BLANK (USPTO)

1945

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2000, 08:19:54 ; Search time 14.43 Seconds
(without alignments)
835.771 Million cell updates/sec

Title: US-08-900-220-17

Perfect score: 2088
Sequence: 1 MALLTNLPICCLALLALPA.....PTGMWYSLLYRLAEELLG 396

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 83857 seqs, 30454973 residues

Total number of hits satisfying chosen parameters: 83857

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : SwissProt_38:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4688	100.0	396	1 DHH_HUMAN	043323 homo sapien
2	2008	96.2	396	1 DHH_MOUSE	061488 mus musculu
3	1273.5	61.0	396	1 DHH1_XENLA	091610 xenopus lae
4	1255.5	60.1	398	1 DHH2_XENLA	091611 xenopus lae
5	1214	58.1	437	1 SHH_MOUSE	062226 mus musculu
6	1203	57.6	437	1 SHH_RAT	063673 rattus norv
7	1198	57.4	425	1 SHH_CHICK	091035 gallus gall
8	1177	56.4	462	1 SHH_HUMAN	013465 homo sapien
9	1161.5	55.6	411	1 THH_HUMAN	014623 homo sapien
10	1160	55.6	408	1 THH_MOUSE	098938 gallus gall
11	1154	55.3	411	1 THH_CHICK	097812 mus musculu
12	1154	55.3	416	1 THH_MOUSE	090419 brachydanio
13	1131.5	54.2	432	1 SHH_MOUSE	090385 cynops pyr
14	1120	53.6	418	1 SHH_MOUSE	092008 brachydanio
15	1092.5	52.3	444	1 SHH_MOUSE	092000 xenopus lae
16	1080.5	51.7	409	1 SHH_MOUSE	091612 xenopus lae
17	1022	48.9	412	1 SHH_MOUSE	098862 brachydanio
18	943	45.2	471	1 HH_MOUSE	002936 drosophila
19	883.5	42.3	481	1 HH_MOUSE	056574 drosophila
20	484.5	23.2	121	1 SHH_MOUSE	076691 carassius a
21	484.5	23.2	121	1 SHH_MOUSE	076691 carassius a
22	482.5	23.1	121	1 SHH_MOUSE	013235 danio aff.
23	482.5	23.1	121	1 SHH_MOUSE	013235 danio aff.
24	482.5	23.1	121	1 SHH_MOUSE	013238 danio aff.
25	482.5	23.1	121	1 SHH_MOUSE	013245 danio aff.
26	482.5	23.1	121	1 SHH_MOUSE	076709 danio kerri
27	482.5	23.1	121	1 SHH_MOUSE	076717 danio pulch
28	482.5	23.1	121	1 SHH_MOUSE	013241 devario dev
29	482.5	23.1	121	1 SHH_MOUSE	013247 devario mal
30	482.5	23.1	121	1 SHH_MOUSE	013247 devario mal
31	482.5	23.1	121	1 SHH_MOUSE	076915 taniichthys
32	482.5	23.1	121	1 SHH_MOUSE	076915 taniichthys
33	480.5	23.0	121	1 SHH_MOUSE	076858 rasbora ele
34	478.5	22.9	121	1 SHH_MOUSE	076864 rasbora het

35	475.5	22.8	121	1 SHH_MOUSE	076662 amblypharyn
36	475.5	22.8	121	1 SHH_MOUSE	076838 punctus con
37	291	13.9	58	1 DHH_MOUSE	076729 brachydanio
38	290	13.9	58	1 DHH_MOUSE	076729 brachydanio
39	290	13.9	58	1 DHH_MOUSE	076693 carassius a
40	290	13.9	58	1 DHH_MOUSE	013240 danio aff.
41	290	13.9	58	1 DHH_MOUSE	076711 danio kerri
42	290	13.9	58	1 DHH_MOUSE	076719 danio pulch
43	290	13.9	58	1 DHH_MOUSE	013243 devario dev
44	290	13.9	58	1 DHH_MOUSE	076852 punctus tet
45	290	13.9	58	1 DHH_MOUSE	076860 rasbora ele

ALIGNMENTS

RESULT 1

ID	Query Match	STANDARD	PRT	396 AA
DHH_HUMAN	043323; Q15794;			
AC	15-JUL-1999 (Rel. 38, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DE	15-JUL-1999 (Rel. 38, Last annotation update)			
GN	DESERET HEDGEHOG PROTEIN PRECURSOR (DHH) (HhG-3).			
OS	Homo sapiens (Human).			
CC	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;			
CC	Eutheria; Primates; Catarrhini; Homiidae; Homo.			
CC	[1]			
RP	SEQUENCE FROM N.A.			
RA	Tate G., Endo Y., Mitsuura T.;			
RL	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.			
RM	[2]			
RP	SEQUENCE OF 85-178 FROM N.A.			
RC	TISSUE-KIDNEY;			
RA	Drummond I.A.;			
RT	"Human desert hedgehog.";			
RL	Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.			
CC	- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT. MAY FUNCTION AS A SPERMATOCYTE SURVIVAL FACTOR IN THE TESTES. ESSENTIAL FOR TESTES DEVELOPMENT.			
CC	- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).			
CC	- PRT: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).			
CC	- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.			
CC	- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AB010994; BAA24866.1;			
DR	EMBL; AB010581; BAA24866.1; JOINED.			
DR	EMBL; AB010993; BAA24866.1; JOINED.			
DR	EMBL; U59748; AAB03398.1;			
DR	PFAM; PF01079; Hhnt; 1.			
DR	PFAM; PF01085; Hh-signal; 1.			
DR	PRINTS; PR00632; SONICHOG.			
KW	Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;			

```

KW  Signal.
FT  SIGNAL. 1 22
FT  CHAIN 23 396
FT  CHAIN 23 396
FT  CHAIN 198 198
FT  CHAIN 199 396
FT  SITE 198 396
FT  SITE 244 244
FT  SITE 244 244
FT  SITE 268 268
FT  ACT_SITE 271 271
FT  BINDING 198 198
FT  CONFLICT 177 177
SQ  SEQUENCE 396 AA: 43577 MW: FCE4FB21972C3AD5 CRC64:

Query Match 100.0% Score 2088: DB 1: Length 396:
Best Local Similarity 100.0% Pred. No. 1.9e-163;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALTLNLPCLALALPAQSCGPGRRYARKQVPLLYKQFVGPVPTLGAS 60
    |||
DB 1 MALTLNLPCLALALPAQSCGPGRRYARKQVPLLYKQFVGPVPTLGAS 60

QY 61 GPAEGRVARSERFDLVNPNYNDIIFKDENSGADRLMTERCKERYNALAIYVNMMPG 120
    |||
DB 61 GPAEGRVARSERFDLVNPNYNDIIFKDENSGADRLMTERCKERYNALAIYVNMMPG 120

QY 121 VRLRVTEGMDHGHADSLHYEGRALDITTSRDRNRYGLLARLAVEGFDWYYESRN 180
    |||
DB 121 VRLRVTEGMDHGHADSLHYEGRALDITTSRDRNRYGLLARLAVEGFDWYYESRN 180

QY 181 HVAHVYKADNSLAVRAGGCEPFGNATVRLMSGERKGLRELHSGDMLAADA5GHNVTPTVL 240
    |||
DB 181 HVAHVYKADNSLAVRAGGCEPFGNATVRLMSGERKGLRELHSGDMLAADA5GHNVTPTVL 240

QY 241 LFLDRDLQRRASVAVETMPRPKLLTPHNLVFAARGPARADPAFARLARGDSV 300
    |||
DB 241 LFLDRDLQRRASVAVETMPRPKLLTPHNLVFAARGPARADPAFARLARGDSV 300

QY 301 LAGGDLARPARAVARAREVAPPLTNGHTLVNDVLA5CYAVLESQVHARAPAPL 360
    |||
DB 301 LAGGDLARPARAVARAREVAPPLTNGHTLVNDVLA5CYAVLESQVHARAPAPL 360

QY 361 RLHLALGALLPGGAVOPTGMWYSRLLYRLAEELLG 396
    |||
DB 361 RLHLALGALLPGGAVOPTGMWYSRLLYRLAEELLG 396

RESULT 2
DHH_MOUSE
ID DHH_MOUSE STANDARD: PRT: 396 AA.
AC 061488:
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE DESERT HEDGEHOG PROTEIN PRECURSOR (DHH) (HMG-3).
GN DHH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN 11)
RP SEQUENCE FROM N.A.
RX STRAIN:129/SV;
RX MEDLINE: 94094334.
RA Echelard Y., Epstein D.J., St Jacques B., Shen L., Mohler J.,
   McMahon J.A., McMahon A.P.;
   "Sonic hedgehog, a member of a family of putative signaling
   molecules, is implicated in the regulation of CNS polarity.";
   Cell 75:1417-1430(1993).
RT 12)
RP SEQUENCE OF 120-168 FROM N.A.

```

```

RX MEDLINE: 95236997.
RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,
RA Zhao R., Seldin M.F., Fallon J.F., Beachy P.A.;
   "Products, genetic linkage and limb patterning activity of a murine
   hedgehog gene.";
   Development 120:339-353(1994).
CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
   PATTERNING EVENTS DURING DEVELOPMENT. MAY FUNCTION AS A
   SPERMATOCTYE SURVIVAL FACTOR IN THE TESTES. ESSENTIAL FOR TESTES
   DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
   CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
   TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ADULT TESTES. NOT EXPRESSED IN
   LIMB BUDS.
CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
   AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
   THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
   A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
   TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
   TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
   OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
   ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
   C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC DR EMBL: X76292; CAA53924.1; -.
CC DR MFAM: PF01079; HInt: 1.
CC DR PFAM: PF01085; HH_signal: 1.
CC DR PRINTS: PR00632; SONICHHOG.
CC KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
   Signal.
KW  Signal.
FT  SIGNAL. 1 22
FT  CHAIN 23 396
FT  CHAIN 23 396
FT  CHAIN 198 198
FT  CHAIN 199 396
FT  SITE 198 198
FT  SITE 244 244
FT  SITE 244 244
FT  SITE 268 268
FT  ACT_SITE 271 271
FT  BINDING 198 198
FT  CONFLICT 177 177
SQ  SEQUENCE 396 AA: 43542 MW: AFEB051BE950FD8 CRC64:

Query Match 96.2% Score 2088: DB 1: Length 396:
Best Local Similarity 96.5% Pred. No. 1.1e-162;
Matches 382; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 MALTLNLPCLALALPAQSCGPGRRYARKQVPLLYKQFVGPVPTLGAS 60
    |||
DB 1 MALTLNLPCLALALPAQSCGPGRRYARKQVPLLYKQFVGPVPTLGAS 60

QY 61 GPAEGRVARSERFDLVNPNYNDIIFKDENSGADRLMTERCKERYNALAIYVNMMPG 120
    |||
DB 61 GPAEGRVARSERFDLVNPNYNDIIFKDENSGADRLMTERCKERYNALAIYVNMMPG 120

QY 121 VRLRVTEGMDHGHADSLHYEGRALDITTSRDRNRYGLLARLAVEGFDWYYESRN 180
    |||
DB 121 VRLRVTEGMDHGHADSLHYEGRALDITTSRDRNRYGLLARLAVEGFDWYYESRN 180

```

QY	181	HHVHSVKADNSLAVAGAGCGFCGNAIVRLMSEERKRLREIHGQDWVLADASGRVPTPYL	240
	181	HHVHSVKADNSLAVAGAGCGFCGNAIVRLMSEERKRLREIHGQDWVLADASGRVPTPYL	240
	181	HHVHSVKADNSLAVAGAGCGFCGNAIVRLMSEERKRLREIHGQDWVLADASGRVPTPYL	240
QY	241	LELDRDLQRRASFVAVETEMPPRKLTLTPMHLVFAARGPAPAPGDFAFVARRLRAGDSV	300
	241	LELDRDLQRRASFVAVETEMPPRKLTLTPMHLVFAARGPAPAPGDFAFVARRLRAGDSV	300
	241	LELDRDLQRRASFVAVETEMPPRKLTLTPMHLVFAARGPAPAPGDFAFVARRLRAGDSV	300
QY	301	LAPGDALRPAPARVAREEAVGVAPDLTANGTLVNDVLAACVAVLESHOMAHFAFPL	360
	301	LAPGDALRPAPARVAREEAVGVAPDLTANGTLVNDVLAACVAVLESHOMAHFAFPL	360
	301	LAPGDALRPAPARVAREEAVGVAPDLTANGTLVNDVLAACVAVLESHOMAHFAFPL	360
QY	361	RLHLRGLLPGGAVOPTGMHWYSRLTLRLAEILLG	396
	361	RLHLRGLLPGGAVOPTGMHWYSRLTLRLAEILLG	396
	361	RLHLRGLLPGGAVOPTGMHWYSRLTLRLAEILLG	396
RESULT	3		
DB	DHHL_XENLA	STANDARD:	PRI: 396 AA.
DT	15-JUL-1999	(Rel. 38, Created)	
DT	15-JUL-1999	(Rel. 38, Last sequence update)	
DT	15-JUL-1999	(Rel. 38, Last annotation update)	
DE	DESERT HEDGEHOG PROTEIN PRECURSOR 1 (DHH-1) (CEPHALIC HEDGEHOG PROTEIN) (X-CHH).		
GN	CHH.		
OS	Xenopus laevis (African clawed frog).		
OC	Eurarchonta; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;		
OC	Batrachia; Anura; Mesobatrachia; Piploidea; Pipidae; Xenopodinae;		
OC	Xenopus.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=EMBRYO;		
RX	MEDLINE: 95401852.		
RA	Ekker S.C., McGrew L.L., Lai C.-J., Lee J.J., von Kessler D.P.,		
RA	Moon R.T., Beachy P.A.:		
RT	"Distinct expression and shared activities of members of the hedgehog		
RT	gene family of Xenopus laevis."		
RL	Development 121:2337-2347(1995).		
CC	-1- FUNCTION: SIGNAL INVOLVED IN THE EARLY INDUCTION AND PATTERNING OF		
CC	ANTERODORSAL ECTODERM, NERVOUS SYSTEM AND SOMITES. INDUCES ECTOPIC		
CC	CEMENT GLAND FORMATION IN EMBRYOS.		
CC	-1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE		
CC	CELL, WHILE THE C-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE		
CC	CELL SURFACE. HEDGEHOG PROTEIN IS ALSO SECRETED IN EITHER CLEAVED		
CC	OR UNCLEAVED FORM TO MEDIATE SIGNALING TO OTHER CELLS (BY		
CC	SIMILARITY).		
CC	-1- DEVELOPMENTAL STAGE: DETECTABLE WITHIN THE EARLY GASTROULA. IN THE		
CC	NEURULA EXPRESSION BECOMES RESTRICTED TO ANTERIOR STRUCTURES,		
CC	ENCOMPASSING BOTH NEURAL PLATE AND ENDODERMAL CELLS.		
CC	-1- PMR: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY		
CC	AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN		
CC	THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF		
CC	A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-		
CC	TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS		
CC	TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION		
CC	OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE		
CC	ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE		
CC	C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).		
CC	-1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.		
CC	-----		
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL: U26349; AA85163.1; -		
DR	PFAM: PF01079; Hint: 1.		
DR	PFAM: PF01085; HH_signal: 1.		

DW	PRINTS: PR00632; SONICHOG.
KR	Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
KM	Signal.
FT	SIGNAL 1 22 POTENTIAL.
FT	CHAIN 23 396 DESERT HEDGEHOG PROTEIN 1.
FT	CHAIN 23 197 DESERT HEDGEHOG PROTEIN 1 N-PRODUCT.
FT	CHAIN 198 396 DESERT HEDGEHOG PROTEIN 1 C-PRODUCT.
FT	SITE 197 198 CLEAVAGE (AUTO-) (BY SIMILARITY).
FT	SITE 267 267 INVOLVED IN AUTO-CLEAVAGE (BY SIMILARITY).
FT	ACT_SITE 270 270 ESSENTIAL FOR AUTO-CLEAVAGE (BY SIMILARITY).
FT	BINDING 197 197 CHOLESTEROL (BY SIMILARITY).
FT	DOMAIN 276 279 POLY-SER.
SD	SEQUENCE 396 AA; 44087 MW; 774A3EC2268A5EE9 CRC64;
<hr/>	
Query Match 61.0%; Score 1273.5; DB 1; Length 396;	
Best Local Similarity 63.6%; Pred. No. 1.7e-108;	
Matches 248; Conservative 52; Mismatches 83; Indels 7; Gaps	
QY	7 LLLPCTALLALPAOSCGPGRGVRRRYARKOLVPLLYKQFYGVPERITAGASGAPGR 66
Dd	8 ILAIIC-GLLIVPYRCGGPGRGVRRRYMRK-LVPLHYGFVPNNPEKITLAGSGNSEKG 65
QY	67 VARGSERPDIVPNVDIIFKDENSGADRLMTERRCKERYNALAIAYNMMPGVRLRV 126
Dd	66 IHRSEFEIEIVPNYNDIIFKDEEKGADRIMTERCKDRYNALATISVMNMPGVKLRT 125
QY	127 EGMDEDHHNODSIHYGRLDITTSRDNRKKGLLARLVAEGFDWVYESNNHVSV 166
Dd	126 EGWEDDHHAHDSIHYGRALDITTSRDNRKKGLMARLVAEGFDWVYESKNHHSV 185
QY	187 KADNSLAVRAGGCPGNAATVRLMSGKERGLRELRGDWVLAADSGRVPTPLYFLDRD 246
Dd	186 KADNSLVGRSGCPRGRAMMMNGEKRPSELKIGDTVTYTDETGLITSVLLFHLRN 245
QY	247 LQRASFVAVEETEMPFRKLLITPMHLVFAARGPARADPAVFARRLAGDSV-IAPGG 305
Dd	246 PYKTATFVLIEAGHPKSLVTTPNHLLFI--QSSSAGFLPTFAVRVQIGDLQIVVNG 302
QY	306 DALPRAVARAKAREANGVAPRLTAHGTLVNDYLASCIAVLESHOWAHAFAPRLIIHA 365
Dd	303 IQVOSSKRVRSLEEQGYVAPMTEHOTLTLDGVLTCYAIVESHILAHS LAPRLIFQG 362
QY	366 LGALLPGGAVOPTGMHWYSRLTYLADEL 395
Dd	363 IASMIP-DLDMSGDVHWYCHILYLAKEYL 391
<hr/>	
RESULT 4	
ID	DHH2_XENLA STANDARD; PRT; 398 AA.
DC	O91611;
DT	15-JUL-1999 (Rel. 38, Created)
DT	15-JUL-1999 (Rel. 38, Last sequence update)
DT	15-JUL-1999 (Rel. 38, Last annotation update)
DE	DESERT HEDGEHOG PROTEIN PRECURSOR 2 (DHH-2) (HEDGEHOG PROTEIN 4) (X-HH4).
GN	HH4.
OS	Xenopus laevis (African clawed frog).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC	Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae; Xenopodinae;
OC	Xenopus.
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=EMBryo;
RX	MEDLINE; 95401852.
RA	Ecker S.C., McGrew L.L., Lai C.-J., Lee J.J., von Kessler D.P.,
RT	"Distinct expression and shared activities of members of the hedgehog
RT	gene family of Xenopus laevis."
RL	Development 121:2337-2347(1995)

```

Db 304 GTOYSSKVVYRVSVDEDTGYVAMNTEHGLVLDVGLTGCATVESHILAHNSLAPLALFQ 363
Oy 365 ALGALDPGAVOPTGMHWYSLRYLAEELL 395
Db 364 GIASMLP-DLHTSDGVHWYCHILYVLAKYVL 393

RESULT 5
SHH_MOUSE STANDARD; PRT; 437 AA.
ID SHH_MOUSE
AC 062226;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE SONIC HEDGEHOG PROTEIN PRECURSOR (SHH) (HhG-1).
GN SHH OR HHG1.
OS Mus musculus (Mouse).
OC Eutheria; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus?
RN
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE; 94094334.
RA Echelard J.A., Epstein D.J., St Jacques B., Shen L., Mohler J.,
RA McMahon J.A., McMahon A.P.;
RT "Sonic hedgehog, a member of a family of putative signaling
RT molecules, is implicated in the regulation of CNS polarity.";
RL Cell 75:1417-1430(1993).
RN
RN REVISION TO 122.
RC STRAIN=C57BL/6J;
RA McMahon A.P.;
RN Submitted (NOV-1997) to the EMBL/Genbank/DBJ databases.
RN
RN [3]
RN SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
RX MEDLINE; 95236997.
RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandi B.K.,
RA Zhao R., Seidlin M.F., Fallon J.F., Beachy P.A.;
RT "Products, genetic linkage and limb patterning activity of a murine
RT hedgehog gene.";
RL Development 120:3339-3353(1994).
RN
RN [4]
RN FUNCTION, AND AUTOPROTEOLYTIC CLEAVAGE.
RX MEDLINE; 95254654.
RA Roelink H., Porter J.A., Chiang C., Tanabe Y., Chang D.T.,
RA Beachy P.A., Jessell T.M.;
RT "Floor plate and motor neuron induction by different concentrations of
RT the amino-terminal cleavage product of sonic hedgehog
RT autoproteolysis.";
RL Cell 81:445-455(1995).
RN
RN [5]
RN X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 34-195.
RX MEDLINE; 96069744.
RA Hall T.M.T., Porter J.A., Beachy P.A., Leahy D.J.;
RT "A potential catalytic site revealed by the 1.7-A crystal structure of
RT the amino-terminal signalling domain of Sonic hedgehog";
RL Nature 378:212-216(1995).
RN
RN -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN
RN ASSOCIATION WITH SMOOTHENED (SMO). TO ACTIVATE THE TRANSCRIPTION
RN OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRRESSSES THE
RN CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER
RN TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A
RN VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED
RN BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE
RN AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE
RN ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH
RN FLOOR PLATE-AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD
RN CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS
RN 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY
RN STIMULATY).
RN
RN -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
RN CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
RN CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM

```

CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).

CC -1- TISSUE SPECIFICITY: EXPRESSED IN A NUMBER OF EMBRYONIC TISSUES

CC INCLUDING THE NOTOCHORD, VENTRAL NEURAL TUBE, FLOOR PLATE, LONG

CC BUD, ZONE OF POLARIZING ACTIVITY AND POSTERIOR DISTAL MESENCHYME

CC OF LIMBS. IN THE ADULT, EXPRESSED IN LUNG AND NEURAL RETINA.

CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE DURING GASTRULATION.

CC -1- INDUCTION: BY RETINOIC ACID.

CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY

CC IN A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN

CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF

CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-

CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS

CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION

CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE

CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE

CC C-PRODUCT HAS NO SIGNALING ACTIVITY.

CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC DR EMBL; X76290; CAA53922.1; -

CC DR PDB; 1VH; 29-JAN-96.

CC DR MGD; MGI:98297; SHH.

CC DR PRAM; PR01079; Hlt; 1.

CC DR PRAM; PR01085; HH_signal; 1.

CC DR PRAM; PR00632; SONICHOG.

CC DR PRAM; PR00632; SONICHOG.

CC Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;

CC Signal; 3D-structure.

CC KW SIGNAL; 1 24

CC FT CHAIN 1 24

CC FT CHAIN 25 437

CC FT CHAIN 25 198

CC FT CHAIN 199 437

CC FT SITE 198 199

CC FT SITE 244 244

CC FT SITE 244 244

CC FT SITE 268 268

CC FT ACT SITE 271 271

CC FT BINDING 198 198

CC FT DOMAIN 383 387

CC SQ SEQUENCE 437 AA; 47773 MW; DOEB72F08E7860EF CRC64;

Query Match 58.1%; Score 1214; DB 1; Length 437;

Best Local Similarity 58.5%; Pred. No. 2e-95;

Matches 248; Conservative 54; Mismatches 82; Indels 40; Gaps 12;

QY 7 LRLPCCIALA-----LPAOSCGPGRGRRRRYARQVLPYLYKQVGPVPTLGAAG 61

DB 4 LRAECFVILIASLLVCPGLACGPRG-FGRRRHPRK-LTPFLAVKQFIPNAEKTILGASG 61

QY 62 PAEGRAVAGSRERFDLPVNPDIIFDEENSGADRLTECKERVNALAVANMMPGV 121

DB 62 RYBCKITRNSRERKELPNINPDILFDEENTGADRLTORCKDKLNALAVSNOMPGV 121

QY 122 RLRTGEMDEDEGHAADSLAHEGRALDITTSRDRNKYGLARLAVLEGFDMVYESRHH 181

DB 122 KLRTGEMDEDEGHSSESLAHEGRAVDITTSRDRSKYGMRLARLAVLEGFDMVYESRAH 181

QY 182 VHSYKADNSLAVAGCGPENGATVRLMSGERKGLRELHRCGDWVLADASRVVPTVLL 241

DB 182 IHCYVKKAKNSVAAKSGGFFGSATVHLEOGCTKLVKDLRPDRVLAADDGRLYSDFLL 241

QY 242 FLDDDLRRASVAVETEMPRKLLTPMHLVFPA-----RGAPAPAGDFAVFAVRRLAG 297

DB 242 FLDDDEGAKKVFYVETLEPRERLLTPAHLFLFVAPHNDSGPTGP--SALFASRVAPG 298

QY 298 DSVIA---PGCD-ALRPARYARYA-REAVGVFAPLTAHGTLLVNDVLAACYAVLESQW 352

DB 299 QRYVVAERGDGDRLLPAAVHSTLREEGAGAYAPLTAHGTLLINRVLAACYAVLESQW 358

QY 353 AHRAFAFRLRLH-LGALLP-----GGAIV-----QPT-GMHWYSRLYLRLA 391

DB 359 AHRAFAFRLRLHALLALAPARDGGGSGIPAAQATARGAEPAGIMYSOLLYHIG 418

QY 392 EELL 395

DB 419 TWLL 422

RESULT 6

ID SHH_RAT STANDARD; PRT; 437 AA.

AC Q63673;

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE SONIC HEDGEHOG PROTEIN PRECURSOR (SHH).

GN SHH OR VHH-1.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-SPRAGUE-DAWLEY, TISSUE-EMBRYONIC FLOOR PLATE;

RX MEDLINE; 94170375.

RA Roelink H., Augsburg A., Heemskerk J., Korzh V., Norlin S.,

RA Ruitz I., Altada A., Tanabe Y., Placzek M., Edlund T., Jessell T.M.,

RA Dodd J.;

RT "Floor plate and motor neuron induction by vhh-1, a vertebrate

RT homolog of hedgehog expressed by the notochord";

RT Cell 76:761-775(1994).

CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN

CC ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION

CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE

CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER

CC TARGET, THE G11 ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A

CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED

CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE

CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE

CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH

CC FLOOR PLATE AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD

CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS

CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY

CC SIMILARITY).

CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE

CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE

CC CELL SURFACE. IT IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM

CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).

CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE NODE, NOTOCHORD, FLOOR PLATE,

CC AND POSTERIOR LIMB BUD MESENCHYME.

CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY

CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN

CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF

CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-

CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS

CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION

CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE

CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE

CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

FT ACT_SITE 273 273 ESSENTIAL FOR AUTO-CLEAVAGE (BY
 FT BINDING 200 200 SIMILARITY).
 FT DOMAIN 390 393 CHOLESTEROL (BY SIMILARITY).
 FT SEQUENCE 425 AA; 46474 MW; DA9627443DA0173 CRC64;
 Query Match 57.4%: Score 1198; DB 1; Length 425;
 Best Local Similarity 59.4%; Pred. No. 4.4e-94;
 Matches 244; Conservative 56; Mismatches 91; Indels 20; Gaps 10;
 QY 1 MALLTNLP---LCCLMLALPAOSCGRGPRVRRARARQVPLLYKQVPPVPEPTL 57
 DB 4 MLLTRILVGFIC--ALVSSGLTCGGNG-ISKRRPKK-LTPPLAKQEPFNPVAKTLL 59
 QY 58 GASPAAGRVARGSEFPDLPVNPPIIFDEENSGADRMTCEKRVNLAIVANNM 117
 DB 60 GASGRYBGKIRNSERKRELTPNNPPIIFDEENSGADRMTCEKRVNLAIVANNM 119
 QY 118 WPGVRLNTEGMDGDHHAODSLHYEGRALDITTSDRDNRYGLLARLAVEGFDWYVE 177
 DB 120 WPGVRLNTEGMDGDHHAODSLHYEGRALDITTSDRDNRYGLLARLAVEGFDWYVE 179
 QY 178 SRNHHVSVYADNLAIVRAGCGFPGNATVRYMSEGRKRLRLHGDVYLVADASGRVPT 237
 DB 180 SKAHHSVYKVENSVAAKSGGCGFPGSAVHLHGGTKLVKDLSPDRVLAADAGRLYS 239
 QY 238 PVLFLDRLDRASFVAVETEMPPKLLTPMHLVFA--RGPAAPAGDPA--PVEARR 293
 DB 240 DFLFLDRMSSKRLFYIETROPARRALLTAHLFLFAAPHNOSSEATGSGALFPMN 299
 QY 294 LRAGDS--VIAFGDALPAPAVANVA--REAVGVAPFLAAGTLVNDVLASCYAVLESH 350
 DB 300 VKPQRYVYVLEGGGQQLLPASVHSVLSREASGAVAPFLAAGTLVNDVLASCYAVIEH 359
 QY 351 QMARAPAPRLRLHA-LGALLPGAV-----QPTGMHYSLLYRLABELL 395
 DB 360 SWAHMAFPRLAAGLLAALCPDAIPTAATTGTHIHYSLLYRIGSWL 410
 RESULT 8
 SHH_HUMAN STANDARD; PRT; 462 AA.
 AC Q15465;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE SONIC HEDGEHOG PROTEIN PRECURSOR (SHH) (HNG-1).
 GN SHH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE: FETAL LUNG;
 RX MEDLINE; 96070431.
 RA Marigo V., Roberts D., Lee S.M.K., Tsukurov O., Levi T.,
 RA Gaslier J.M., Epstein D.J., Gilbert D.J., Copeland N.G., Seidman C.E.,
 RA Jenkins N.A., Seidman J.G., McMahon A.P., Tabin C.;
 RT "Cloning, expression, and chromosomal location of SHH and IHH: two
 RT human homologues of the Drosophila segment polarity gene hedgehog.";
 RL Genomics 28:44-51(1995).
 RN [2]
 RP SEQUENCE OF 1-187 FROM N.A.
 RA Strong C., Graves T., Suterer C., Ozersky P.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 119-167 FROM N.A.
 RX MEDLINE; 95236997.
 RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,
 RA Zhao R., Seidlin M.F., Fallon J.F., Beachy P.A.;
 RT "Products, genetic linkage and limb patterning activity of a murine
 hedgehog gene.";

RL Development 120:3339-3353(1994).
 RN [4]
 RP VARIANTS HPE3 ARG-31; GLY-117 AND ARG-117.
 RX MEDLINE; 97051937.
 RA Roessler E., Belloni E., Gaudenz K., Jay P., Berta P., Scherer S.W.,
 RA Tsui L.-C., Muenke M.;
 RT "Mutations in the human Sonic Hedgehog gene cause holoprosencephaly.";
 RL Nat. Genet. 14:357-360(1996).
 CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN
 CC ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION
 CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE
 CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER
 CC TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A
 CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED
 CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE
 CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE
 CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH
 CC FLOOR PLATE AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD
 CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS
 CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL. WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN FETAL INTESTINE, LIVER, LUNG, AND
 CC KIDNEY. NOT EXPRESSED IN ADULT TISSUES.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- DISEASE: DEFECTS IN SHH ARE THE CAUSE OF THE AUTOSOMAL DOMINANT
 CC DISORDER HOLOPROSENCEPHALY TYPE 3 (HPE3). HPE3 IS A GENETICALLY
 CC HETEROGENEOUS DISEASE THAT AFFECTS THE MIDLINE DEVELOPMENT OF THE
 CC FOREBRAIN AND MIDFACE. HPE IS ASSOCIATED WITH SEVERAL DISTINCT
 CC FACIES AND PHENOTYPIC VARIABILITY. IN THE MOST EXTREME CASES,
 CC ANOPHTHALMIA OR CYCLOPIA IS EVIDENT ALONG WITH A CONGENITAL
 CC ABSENCE OF THE MATURE NOSE. THE LESS SEVERE FORM FEATURES FACIAL
 CC DYSMORPHIA CHARACTERIZED BY OCULAR HYPERTELEORISM, DEFECTS OF THE
 CC UPPER LIP AND/OR NOSE, AND ABSENCE OF THE OLFACTORY NERVES OR
 CC CORPUS CALLOSUM.
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L38518; AAA62179.1; -
 DR EMBL; AC002484; AAB67604.1; -
 DR MIM; 600725; -
 DR MIM; 142945; -
 DR PFM; PFO1079; Hnt; 1.
 DR PFM; PFO1085; HH_signal; 1.
 DR PRINTS; PRO0632; SONICHOG.
 KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 KW Signal; Disease mutation.
 FT SIGNAL 1 23
 FT CHAIN 24 462
 FT CHAIN 24 197
 FT CHAIN 198 462
 FT SITE 197 198
 FT SITE 243 243
 FT SITE 267 267
 FT SITE 267 267
 INVOLVED IN AUTO-CLEAVAGE (BY

FT	ACT SITE	270	270	SIMILARITY).
FT	BINDING	197	197	ESSENTIAL FOR AUTO-CLEAVAGE (BY
FT	DOMAIN	407	411	SIMILARITY).
FT	VARIANT	31	31	CHOLESTEROL (BY SIMILARITY).
FT	VARIANT	117	117	POLY-GLY
FT	VARIANT	117	117	G -> R (IN HPE3).
FT	VARIANT	117	117	/FTID-VAR_003619.
FT	VARIANT	117	117	W -> G (IN HPE3).
FT	VARIANT	117	117	/FTID-VAR_003620.
FT	VARIANT	117	117	W -> R (IN HPE3).
FT	VARIANT	117	117	/FTID-VAR_003621.
FT	VARIANT	117	117	/FTID-VAR_003622.
FT	VARIANT	117	117	/FTID-VAR_003623.
FT	VARIANT	117	117	/FTID-VAR_003624.
FT	VARIANT	117	117	/FTID-VAR_003625.
FT	VARIANT	117	117	/FTID-VAR_003626.
FT	VARIANT	117	117	/FTID-VAR_003627.
FT	VARIANT	117	117	/FTID-VAR_003628.
FT	VARIANT	117	117	/FTID-VAR_003629.
FT	VARIANT	117	117	/FTID-VAR_003630.
FT	VARIANT	117	117	/FTID-VAR_003631.
FT	VARIANT	117	117	/FTID-VAR_003632.
FT	VARIANT	117	117	/FTID-VAR_003633.
FT	VARIANT	117	117	/FTID-VAR_003634.
FT	VARIANT	117	117	/FTID-VAR_003635.
FT	VARIANT	117	117	/FTID-VAR_003636.
FT	VARIANT	117	117	/FTID-VAR_003637.
FT	VARIANT	117	117	/FTID-VAR_003638.
FT	VARIANT	117	117	/FTID-VAR_003639.
FT	VARIANT	117	117	/FTID-VAR_003640.
FT	VARIANT	117	117	/FTID-VAR_003641.
FT	VARIANT	117	117	/FTID-VAR_003642.
FT	VARIANT	117	117	/FTID-VAR_003643.
FT	VARIANT	117	117	/FTID-VAR_003644.
FT	VARIANT	117	117	/FTID-VAR_003645.
FT	VARIANT	117	117	/FTID-VAR_003646.
FT	VARIANT	117	117	/FTID-VAR_003647.
FT	VARIANT	117	117	/FTID-VAR_003648.
FT	VARIANT	117	117	/FTID-VAR_003649.
FT	VARIANT	117	117	/FTID-VAR_003650.
FT	VARIANT	117	117	/FTID-VAR_003651.
FT	VARIANT	117	117	/FTID-VAR_003652.
FT	VARIANT	117	117	/FTID-VAR_003653.
FT	VARIANT	117	117	/FTID-VAR_003654.
FT	VARIANT	117	117	/FTID-VAR_003655.
FT	VARIANT	117	117	/FTID-VAR_003656.
FT	VARIANT	117	117	/FTID-VAR_003657.
FT	VARIANT	117	117	/FTID-VAR_003658.
FT	VARIANT	117	117	/FTID-VAR_003659.
FT	VARIANT	117	117	/FTID-VAR_003660.
FT	VARIANT	117	117	/FTID-VAR_003661.
FT	VARIANT	117	117	/FTID-VAR_003662.
FT	VARIANT	117	117	/FTID-VAR_003663.
FT	VARIANT	117	117	/FTID-VAR_003664.
FT	VARIANT	117	117	/FTID-VAR_003665.
FT	VARIANT	117	117	/FTID-VAR_003666.
FT	VARIANT	117	117	/FTID-VAR_003667.
FT	VARIANT	117	117	/FTID-VAR_003668.
FT	VARIANT	117	117	/FTID-VAR_003669.
FT	VARIANT	117	117	/FTID-VAR_003670.
FT	VARIANT	117	117	/FTID-VAR_003671.
FT	VARIANT	117	117	/FTID-VAR_003672.
FT	VARIANT	117	117	/FTID-VAR_003673.
FT	VARIANT	117	117	/FTID-VAR_003674.
FT	VARIANT	117	117	/FTID-VAR_003675.
FT	VARIANT	117	117	/FTID-VAR_003676.
FT	VARIANT	117	117	/FTID-VAR_003677.
FT	VARIANT	117	117	/FTID-VAR_003678.
FT	VARIANT	117	117	/FTID-VAR_003679.
FT	VARIANT	117	117	/FTID-VAR_003680.
FT	VARIANT	117	117	/FTID-VAR_003681.
FT	VARIANT	117	117	/FTID-VAR_003682.
FT	VARIANT	117	117	/FTID-VAR_003683.
FT	VARIANT	117	117	/FTID-VAR_003684.
FT	VARIANT	117	117	/FTID-VAR_003685.
FT	VARIANT	117	117	/FTID-VAR_003686.
FT	VARIANT	117	117	/FTID-VAR_003687.
FT	VARIANT	117	117	/FTID-VAR_003688.
FT	VARIANT	117	117	/FTID-VAR_003689.
FT	VARIANT	117	117	/FTID-VAR_003690.
FT	VARIANT	117	117	/FTID-VAR_003691.
FT	VARIANT	117	117	/FTID-VAR_003692.

GN ITH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RN SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
RC STRAIN-CD-1; TISSUE-KIDNEY;
RX MEDLINE: 97236802.
RA Valentini R.P., Brookhiser W.T., Park J., Yang T., Briggs J.,
RA Dressler G., Holzman L.B.;
RA "Post-translational processing and renal expression of mouse Indian
RT hedgehog".
RN J. Biol. Chem. 272:8466-8473(1997).
RN [2]
RN SEQUENCE OF 76-411 FROM N.A.
RC STRAIN-C57BL/6J; TISSUE-EMBRYO;
RX MEDLINE: 94094334.
RA Echelard Y., Epstein D.J., St Jacques B., Shen L., Mohler J.,
RA McMahon J.A., McMahon A.P.;
RA "Sonic hedgehog, a member of a family of putative signaling
RT molecules, is implicated in the regulation of CNS polarity".
RN Cell 75:1417-1430(1993).
RN [3]
RN REVISIONS.
RP STRAIN-C57BL/6J;
RA St Jacques B.;
RN Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RN SEQUENCE OF 124-172 FROM N.A.
RX MEDLINE: 95236997.
RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,
RA Zhao R., Seidlin M.F., Fallon J.F., Beachy P.A.;
RA "Products, genetic linkage and limb patterning activity of a murine
RT hedgehog gene".
RN Development 120:3339-3353(1994).
RN [1]
RN FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
CC PATTERNING EVENTS DURING DEVELOPMENT. BINDS TO THE PATCHED (PTC)
CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHED (SMO), TO
CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IMPLICATED IN
CC ENDODERMAL OSSIFICATION: MAY REGULATE THE BALANCE BETWEEN GROWTH
CC AND OSSIFICATION OF THE DEVELOPING BONES. INDUCES THE EXPRESSION
CC OF PARATHYROID HORMONE-RELATED PROTEIN (PTHrP)
CC OF PARATHYROID LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC TISSUE SPECIFICITY: IN THE EMBRYO, DETECTED IN THE DEVELOPING GUT,
CC THE GROWTH ZONE OF CARTILAGE OF DEVELOPING LONG BONES, EPITHELIUM
CC AND UROGENITAL SINUS. IN THE ADULT KIDNEY, FOUND IN PROXIMAL
CC CONVOLUTED AND PROXIMAL STRAIGHT TUBULE.
CC [1]
CC DEVELOPMENTAL STAGE: DETECTED AT 10 DAYS POST COITUM (DPC) IN
CC DEVELOPING GUT, AT 14.5 DAYS DPC IN THE CARTILAGE PRIMORDIUM AND
CC IN THE DEVELOPING UROGENITAL SINUS. EXPRESSION INCREASES WITH
CC GESTATIONAL AGE IN KIDNEY AND DUODENUM, BECOMING MAXIMAL IN
CC ADULTHOOD.
CC [1]
CC PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC [1]
CC SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
CC [1]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL: U85610; AAB49692.1; ALT_INIT.
DR EMBL: X76291; CAA53923.1; -
DR MGI: 96533; ITH.
DR PFAM: PF01079; Hint: 1.
DR PFAM: PF01085; HH_Signal: 1.
DR PRINTS: PR00632; SONCHHOG.
KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
KW Signal.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 411 INDIAN HEDGEHOG PROTEIN.
FT CHAIN 28 202 INDIAN HEDGEHOG PROTEIN N-PRODUCT.
FT SITE 203 411 INDIAN HEDGEHOG PROTEIN C-PRODUCT.
FT SITE 202 203 CLEAVAGE (AUTO-).
FT SITE 248 248 INVOLVED IN CHOLESTEROL TRANSFER (BY
FT SITE 248 248 SIMILARITY).
FT SITE 272 272 INVOLVED IN AUTO-CLEAVAGE (BY
FT ACT_SITE 275 275 SIMILARITY).
FT BINDING 202 202 ESSENTIAL FOR AUTO-CLEAVAGE (BY
FT CONFLICT 383 383 CHOLESTEROL (BY SIMILARITY).
FT FT 383 W -> S (IN REF. 2).
SQ SEQUENCE 411 AA: 45485 MM: 08BE7AD8507C0D9B CRC64.

Query Match 55.3%; Score 1154; DB 1; Length 411;
Best Local Similarity 59.0%; Pred. No. 2,3e-90;
Matches 230; Conservative 49; Mismatches 99; Indels 12; Gaps 4;

QY 12 CLALLAL---PAQSGGPGGVGRRRTARKQVLPILYKQFVGPVPTLGASGPAEGV 67
DB 13 CFFLLLLVLPARGGPGGR-VVGSRRRPRKVLPLAYKQFPPNPEKTLGASGREGRT 71
QY 68 ARGSEFRDLVNNYNDIFKDEENSGADRLMTGKKEKNALATAYMMMPGVRLE 127
DB 72 ANSEFFKLTPYNDLIFKDEENSGADRLMTGKKEKNALATAYMMMPGVRLE 131
QY 128 GWDEGHHADSLHYEGRALDITTSDBRNNKGLLARLAVEAGFWMVYSENNHVSVK 187
DB 132 GWDEGHHSESLHYEGRAVDITTSDBRNNKGLLARLAVEAGFWMVYSENNHVSVK 191
QY 188 ADNSLAVRAGGCGPGGATVRLMSGKGLRELRGDDVLAALASGRVPTPLFLDR 247
DB 192 SESSAAKTKGCGFPAGQVRLNGSERVALSAVPGGVLAEMGEDGPTFSDVLI 251
QY 248 ORRASVAVETMPRPKLLTPMHLVFAARAPAPGDPAPFARLRAGDSVLAPGDA 307
DB 252 NLRARQVETQDPRLRLTPAHLFTADNHTERPAHFRATFASHVQGYVLVSGV 311
QY 308 LRPARYARAREAVGVFAPLTAHGLTLVNDVLAACYAVLESHOMAHRAFPALRL 367
DB 312 LQPARAVASTHVALGSAVPLTRHGLVEDVVAASFAVADHLLAQALFWPLRL 371
QY 367 -GALPGAVOPTGMHYSRLYLAEELL 395
DB 372 WGSWTPS-----EGVHWYQMLRGLRL 396

RESULT 12
TWHH_BRARE
ID TWHH_BRARE STANDARD; PRT; 416 AA.
AC Q90419;
DT 15-JUL-1999 (rel. 38, Created)
DT 15-JUL-1999 (rel. 38, Last sequence update)
DT 15-FEB-2000 (rel. 39, Last annotation update)
DE TIGGY-WINKLE HEDGEHOG PROTEIN PRECURSOR (TWHH).
GN TWHH.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinidae; Rasbora; Rasbora; Dario.
RN [1]
RN SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.

PFAM; PF01079; Hint: 1.
 DR PFAM; PF01085; Hh_signal: 1.
 DR PRINTS; PR00632; SONICHHOG.
 KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 Signal.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 432 SONIC HEDGEHOG PROTEIN.
 FT CHAIN 27 200 SONIC HEDGEHOG PROTEIN N-PRODUCT.
 FT CHAIN 201 432 SONIC HEDGEHOG PROTEIN C-PRODUCT.
 FT SITE 200 201 CLEAVAGE (AUTO-) (BY SIMILARITY).
 FT SITE 268 268 INVOLVED IN AUTO-CLEAVAGE (BY
 SIMILARITY).
 FT ACT_SITE 271 271 ESSENTIAL FOR AUTO-CLEAVAGE (BY
 SIMILARITY).
 FT BINDING 200 200 CHOLESTEROL (BY SIMILARITY).
 FT SEQUENCE 432 AA; 47847 MW; B455C7E746C8E5A8 CRC64;
 SO
 Query Match 54.2%; Score 1131.5; DB 1; Length 432;
 Best Local Similarity 55.58; Pred. No. 1.9e-88;
 Matches 234; Conservative 57; Mismatches 96; Indels 35; Gaps 9;
 QY 1 MALLINLP---LCCIALALPAQSGPGRGPRRRYARKOLVPLLYKQFVGPVPTL 57
 DB 4 MILLARVLLAGTIC--ALLVPSGLSCGPRGIGTKRF--KLTPLAYKQFTPNVPEKTL 59
 QY 58 GASGPAEGRVANGSERFNDLVNPNPDIIFKDEENSGADRLMTERCKERNALAIAYNM 117
 DB 60 GASGVEGKRTNSERFKELTPNNDIIFKDEENTGADRLMTQCKDKRLNALAIAYNMQ 119
 QY 118 WGVGLRVTEGDEDEGHADODSLAYEGRALDITSDRONKKGKLARLAVENGFWYITE 177
 DB 120 WGVGLRVTEGDEDEGHAEESLHREGRAVDITTDROSKTKGMARLAAEGFWYITE 179
 QY 178 SRNHVAVSKADNSIAYVAGCGCPGNATVRLWSGERKGLRELRHGDWYLAADSGRVPT 237
 DB 180 SKAHHCVCYKAMNSVAAVSGCGFPSSATVLAEGQVRIPIKDLRPDRYLAADGLKLYIS 239
 QY 238 PVLFLDLRLORASFAVETEMPRLKLLPWLHVA--ARGPAPADGFAPV----- 290
 DB 240 DFLFMDEEETVRKYFYVYIETS--REVRVLTAAHLFLFGQAHGNDSGDFSVGSAGF 297
 QY 290 ---FARRRAGDSVAPG--GDALRPAYARAREAGVAPPLAHGTLVNDVLACY 344
 DB 298 RSMFASVYRAGRVLTVDREGRLAEATVERVLEAGVAPVTAHGTIVYDRLACY 357
 QY 345 AVLESHQWAAHRAFAPLRLHALGAL-----LPGAVQPTGMWYSRLYLAEE 393
 DB 358 AVIEHSHMAHNAFAPLRV--GLGILSFTSPQDYSSHPAPSGGVHYSILIRIGTW 415
 QY 394 LL 395
 DB 416 VL 417
 RESULT 14
 SHH_BRARE STANDARD; PRT; 418 AA.
 AC 092008: 013170; 013171;
 DT 15-JUL-1999 (Rel. 38, Last Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE SONIC HEDGEHOG PROTEIN PRECURSOR (SHH) (VHH-1).
 GN SHH OR VHH1.
 OS Brachydanio rerio (zebrafish) (zebra danio).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
 CC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
 CC Cyprinoidae; Cyprinidae; Rasbortinae; Danio.
 CC [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE-EMBRYO;
 RX MEDLINE; 94170375.
 RA Roelink H., Augsburg J., Heemsereck J., Korzh V., Norlin S.,

RA Ruiz I Altaba A., Tanabe Y., Placzek M., Edlund T., Jessell T.M.,
 RA Dodd J.;
 RT "Floor plate and motor neuron induction by vhh-1, a vertebrate
 RT homolog of hedgehog expressed by the notochord.";
 RL Cell 76:761-775(1994).
 RN [2]
 RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
 RX MEDLINE; 96014264.
 RA Ekker S.C., Ungar A.R., Greenstein P., von Kessler D.P., Porter J.A.,
 RA Moon R.T., Beachy P.A.;
 RT "Patterning activities of vertebrate hedgehog proteins in the
 RT developing eye and brain.";
 RL Curr. Biol. 5:944-953(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 96083328.
 RA Fietz M.J., Concordet J.-P., Barbosa R., Johnson R., Krauss S.,
 RA McMahon A.P., Tabin C., Ingham P.W.;
 RT "The hedgehog gene family in Drosophila and vertebrate development.";
 RL Development Suppl. 43-51(1994).
 RN [4]
 RP SEQUENCE OF 30-92 AND 113-170 FROM N.A.
 RC TISSUE-MUSCLE;
 RX MEDLINE; 97075114.
 RA Zardoya R., Abouheif E., Meyer A.;
 RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
 RT closely related to the zebrafish.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE
 CC NOTOCHORD THAT INDUCES SOMITE PATTERNING, DORSO-VENTRAL PATTERNING
 CC OF THE BRAIN AND EARLY PATTERNING OF THE DEVELOPING EYES. DISPLAYS
 CC FLOOR PLATE-INDUCING ACTIVITY. BINDS TO THE PATCHED (PTC)
 CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHEDEN (SMO), TO
 CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH,
 CC PTC REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE VENTRAL MIDLINE OF THE NEURAL
 CC TUBE AND BRAIN. ALSO FOUND IN THE NOTOCHORD AND IN DEVELOPING FIN
 CC BUD. IN THE DEVELOPING BRAIN, EXPRESSION OCCURS IN DOMAINS THAT
 CC INCLUDE A DISCRETE REGION IN THE FLOOR OF THE DIENCEPHALON.
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE IN THE INNER CELL LAYER OF
 CC THE EMBRYONIC SHIELD DURING GASTRULATION. BY 9.5 HR OF
 CC DEVELOPMENT, EXPRESSED IN A CONTINUOUS BAND THAT EXTENDS FROM THE
 CC TAIL TO THE HEAD, THE ANTERIOR BOUNDARY OF EXPRESSION BEING
 CC POSITIONED IN THE CENTER OF THE ANIMAL POLE ANTERIOR TO THE
 CC PRESUMPTIVE MIDBRAIN.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; L27585; AAA20998.1; -
 CC EMBL; U30711; AAC59742.1; -
 CC EMBL; Z35669; CAA84738.1; -

EMBL: U51351; AAB38575.1; -;
 DR EMBL: U51370; AAB38593.1; -;
 DR ZFIN: ZDB-GENE-980526-166; SHH.
 DR PFAM: PF01079; Hint. 1.
 DR PFAM: PF01085; HH_signal; 1.
 DR PRINTS: PRO0632; SONICHOG.
 KM Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 KW Signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 418 SONIC HEDGEHOG PROTEIN.
 FT CHAIN 24 197 SONIC HEDGEHOG PROTEIN N-PRODUCT.
 FT CHAIN 198 418 SONIC HEDGEHOG PROTEIN C-PRODUCT.
 FT SITE 197 198 CLEAVAGE (AUTO-).
 FT SITE 243 243 INVOLVED IN CHOLESTEROL TRANSFER (BY
 FT SITE 267 267 SIMILARITY).
 FT ACT_SITE 270 270 INVOLVED IN AUTO-CLEAVAGE (BY
 FT BINDING 197 197 ESSENTIALITY FOR AUTO-CLEAVAGE (BY
 FT SEQUENCE 418 AA; 46402 MW; CF000AFED2F5795 CRC64;
 SQ
 Query Match 53.68; Score 1120; DB 1; Length 418;
 Best Local Similarity 54.38; Pred. No. 1.7e-87;
 Matches 220; Conservative 66; Mismatches 107; Indels 12; Gaps 5;
 Y 1 MALTLNLPCLCAL-LALPAQSCGPGPGVGRRRYARKOLVPLLYKQVGPVPTLGA 59
 1 MLRLIRLVLSLTLVSLVSLAGCPGPG-YGRRRHPRK-LTFLAYKQFIVNAEKLIGA 58
 Y 60 SGPAERVAQSEERFRLVPIYNDIIFKDENSGADRLTERKERVNALAIAMNMP 119
 59 SGREGKTRSEKFEKELTPYNDIIFKDENTGADRLMORCKDLSIAISVMNMP 118
 Y 120 GYRLVTEGMEDEHHAODSLHTEGRALDITSDRDKRYGLRLAVEAGFDWYYSR 179
 119 GYRLVTEGMEDEHHAODSLHTEGRALDITSDRDKRYGLRLAVEAGFDWYYSR 178
 Y 180 NHVAVYKADSLAVRAGCGFPGNATVRLSGERKGLRELRGDVLAAASGRVPTPV 239
 179 AHICSVKAEVSAKSGCGFPGSALVSLQGGCAKAVDLNPGDKVLAADSNGLVSEDF 238
 Y 240 LFLDRDLQRRASVAVETEMPRLTLTPWHLVFAARGAPADGEPVFAERRLAGDS 299
 239 IMFEDRSTTRVRYVDTQEPVEKILITLAHLFLVDLNSDEDLHTMAAASSVRAQK 298
 Y 300 VL-APGGDALPARYARAEAVGVAPLTAHGLTLVNDVLAACYAVLASHQMAHRAFA 358
 299 VMVVDSSQLSVIVORLYTEORSEFAPVTAHGTIVDRILASCYAVIEDQGLAHLAFA 358
 Y 359 PLRLHLAGALL-----PGAVOPTGMHYSRLVRLABELL 395
 359 PARLYYVSFLFPONSSSRNATLQDGVHMYSLRLQKGTWLL 403
 DB
 RESULT 15
 SHH_XENLA STANDARD; PRT: 444 AA.
 AC Q93000; Q91894;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-FEB-2000 (Rel. 39, Last annotation update)
 DE SONIC HEDGEHOG PROTEIN PRECURSOR (X-SHH) (YHH-1).
 GN SHH.
 OS Xenopus laevis (African clawed frog).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
 CC Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae;
 CC Xenopus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-INTERESTING.
 RX MEDLINE; 95357169.

RA Stelow M.A., Shi Y.-B.;
 RT "Xenopus sonic hedgehog as a potential morphogen during embryogenesis
 RT and thyroid hormone-dependent metamorphosis";
 RL Nucleic Acids Res. 23:2555-2562(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-EMBRYO;
 RX MEDLINE; 95401852.
 RA Erker S.C., McGrew L.L., Lai C.-J., Lee J.J., von Kessler D.P.,
 RA Moon R.T., Beachy P.A.;
 RT "Distinct expression and shared activities of members of the hedgehog
 RT gene family of Xenopus laevis";
 RL Development 121:2357-2347(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-NOTOCHORD;
 RX MEDLINE; 96028338.
 RA Ruiz I Altada A., Jessell T.M., Roelink H.;
 RT "Restrictions to floor plate induction by hedgehog and winged-helix
 RT genes in the neural tube of frog embryos";
 RL Mol. Cell. Neurosci. 6:106-121(1995).
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
 CC PATTERNING EVENTS DURING DEVELOPMENT AND METAMORPHOSIS. INVOLVED
 CC IN LIMB FORMATION. PATTERNING OF THE CENTRAL NERVOUS SYSTEM AND
 CC VENTRAL SOMITE DIFFERENTIATION. INDUCES ECTOPIC CEMENT GLAND
 CC FORMATION IN EMBRYOS. BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH
 CC FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE
 CC TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC
 CC REPRESENTS THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN NOTOCHORD AND NEURAL
 CC FLOOR PLATE DURING EMBRYOGENESIS. IN TADPOLE, HIGH EXPRESSION IS
 CC OBSERVED IN PANGRAFS/STOMACH, MODERATE EXPRESSION IN TAIL, AND LOW
 CC EXPRESSION IN INTESTINE, BRAIN, AND HIND LIMB.
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED AT THE NEURULA (STAGES 16-17).
 CC FIRST PEAK OF EXPRESSION AROUND TADPOLE HATCHING (STAGES 33-40).
 CC HIGH EXPRESSION OBSERVED IN INTESTINE AT THE CLIMAX OF
 CC MORPHOGENESIS (STAGES 60-62) WHEN INTESTINE EPITHELIAL UNDERGOES
 CC MORPHOGENESIS.
 CC -1- INDUCTION: BY THYROID HORMONE.
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L39213; AAC4227.1; -;
 DR EMBL: U26314; AAA85162.1; -;
 DR EMBL: L35248; AAA4981.1; -;
 DR PFAM: PF01079; Hint. 1.
 DR PFAM: PF01085; HH_signal; 1.
 DR PRINTS: PRO0632; SONICHOG.
 KM Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
 KW Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 444 SONIC HEDGEHOG PROTEIN.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2000, 08:19:30 ; Search time 26.41 Seconds
(without alignments)
1039,618 Million cell updates/sec

Title: US-08-900-220-17
Perfect score: 2088
Sequence: 1 MALLTNLPCLCLALLALPA.....PTGMHWYSLRYLAELLG 396

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 225878 segs, 6934122 residues
Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :
1: SP_Archea:*
2: SP_Bacteria:*
3: SP_Fungi:*
4: SP_Human:*
5: SP_Invertebrate:*
6: SP_Mammal:*
7: SP_Mhc:*
8: SP_Organelle:*
9: SP_Phage:*
10: SP_Plant:*
11: SP_Protoct:*
12: SP_Virus:*
13: SP_Vertebrate:*
14: SP_Unclassified:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1144.5	54.8	406	13	057567 notophthalm
2	1126	53.9	414	13	09W709 parolichthy
3	1107.5	53.0	434	13	057404 pleurodeles
4	990	47.4	442	13	073803 fugu rubrip
5	919	44.0	415	5	017499 branchiost
6	781.5	37.4	410	5	061676 lytechinus
7	664	31.8	129	11	09W0P6 rattus norv
8	624	29.9	177	11	09W29 rattus norv
9	618	29.6	139	6	09XSI6 bos taurus
10	606	29.0	150	13	09YGV7 ambystoma m
11	599	28.7	138	13	09W6C1 eleutheroda
12	585	28.0	185	5	096699 junonia coe
13	431	20.6	119	13	042128 oryzias lat
14	375	18.0	88	13	09YGV3 brachydanio
15	352	16.9	80	13	042441 oryzias lat
16	273.5	13.1	137	13	042234 coturnix co
17	209	10.0	868	5	09XV14 caenorhabdi
18	194.5	9.3	790	5	022872 caenorhabdi
19	184.5	9.3	1226	5	021835 caenorhabdi
20	193.5	9.3	1021	5	09XUV2 caenorhabdi

21	186.5	8.9	1207	5	021535 caenorhabdi
22	182	8.7	481	5	045992 caenorhabdi
23	179.5	8.6	557	5	094129 caenorhabdi
24	163	7.8	485	5	094128 caenorhabdi
25	161.5	7.7	550	5	094130 caenorhabdi
26	161.5	7.7	629	5	045273 caenorhabdi
27	150.5	7.2	615	5	091573 caenorhabdi
28	135	6.5	205	5	023193 caenorhabdi
29	115.5	5.5	3670	2	0924X5 streptomyce
30	114.5	5.5	687	2	P71196 escherichia
31	106.5	5.1	481	5	094410 caenorhabdi
32	103.5	5.0	3391	12	092834 dengue viru
33	103.5	5.0	3391	12	092835 dengue viru
34	103	4.9	672	2	09X856 streptomyce
35	102.5	4.9	3391	12	011875 dengue viru
36	102.5	4.9	3391	12	009234 unidentified
37	102.5	4.9	3391	12	092752 dengue viru
38	102.5	4.9	3391	12	092753 dengue viru
39	102.5	4.9	3391	12	09W1Z8 dengue viru
40	102.5	4.9	3391	12	09W1Z7 dengue viru
41	102.5	4.9	3391	12	09W1Z4 dengue viru
42	102.5	4.9	3391	12	09WDA7 dengue viru
43	102.5	4.9	3391	12	09WDA6 dengue viru
44	102.5	4.9	3391	12	09WDA5 dengue viru
45	102.5	4.9	3391	12	09WDA4 dengue viru

ALIGNMENTS

RESULT 1
ID 057567 PRELIMINARY; PRT; 406 AA.
AC 057567;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE HEDGEHOG SEGMENT POLARITY HOMOLOG.
OS Notoththalmus viridescens (Eastern newt) (Triturus viridescens).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC Batrachia; Caudata; Salamandriodea; Salamandridae; Notoththalmus.
RN [1]
RP SEQUENCE FROM N.A.
RA STARK D.R., GATES P.B., BROCKES J.P., FERRETTI P.;
RL Dev. Dyn. 0:0-0(1998).
DR EMBL; AF047466; AAC03108.1; -.
DR HSSP; Q62226; LVH.
DR PFM; PFI01085; HH_signal; 1.
DR PFM; PFI01079; HInt; 1.
DR PRINTS; PR00632; SONICHOG.
SQ SEQUENCE 406 AA; 45072 MW; 9D0FFA76 CRC32;

Query Match 54.8%; Score 1144.5; DB 13; Length 406;
Best Local Similarity 58.2%; Pred. No. 2.1e-81;
Matches 227; Conservative 49; Mismatches 107; Indels 7; Gaps 3;

QY 7 LPLCLALLALP-NQSGPGRGVGRYARKQLVPLTKYQFVPGVPTLIGASPAEG 65
DB 8 LLAVALLLGLGPGALGGPGR--VIGRRPPRLPIPLSTYQFPLPHVEKTLGASGRYEG 65
QY 66 RVANGSERFDLPVNPYNDIFKDENSGADRLMTERCKEYVNLALAVMMAGVRLRV 125
DB 66 KIANNSEFKELTNNYNDIFKDENTGADRLMTQCKDKRLNSLAISVNMQGVKIRV 125
QY 126 TEGWDEGHHADSLHVEGRALDITTSDDRNKXGLLARLAVEAGFPMWYVESNNHVS 185
DB 126 TEGWDEGHHSDSLHVEGRAVDITTSDDRNKXGLLARLAAGFPMWYVESKAHICS 185
QY 186 VKAANSIAVARGCGFPAGATYRLMSGERKGLREHGRDWTYLAASGSRVYPTVLLDR 245
DB 186 VKSHSAAKTGGCFPAALATLESGERKPIADLEPGHRYLCMDGGRRTYSDFLTFLDR 245

QY 246 DLORRASFAVETEMPPKLLTPMHLVFAARGAPAPGDPFAVFAARLACGSVLAPCG 305
DB 246 DSRVAFEEFYVERDRPPRLALTAHLFTVADNFTVPLDFTSFVFSHWOPGQYILTEGV 305
QY 306 DALRPARVAVAREEAVGAPLTAAGTLLVNDVLASCYAVLASHQMAHAFAPRLRLHA 365
DB 306 LGIQPARVAVSVTQTDSGAAPLTLSHOTLLVDVYVSCFAVOKHQAALFAPRLRLXHS 365
QY 366 LGALLPEGAVOPTGMHWYSLLYRLAEELL 395
DB 366 VGR-----PETQPEGMHWSLLYRLKGVLL 391

RESULT 2

Q9W709 PRELIMINARY; PRT; 414 AA.
AC Q9W709; 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DE 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
GN SONIC HEDGEHOG.
OS Paralicthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
OC Pleuronectiformes; Pleuronectoidel; Bothidae; Paralicthys.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 99238226.
RA SUZUKI T., ICHIRO O., KUROKAWA T.;
RT "Retinoic acid given at late embryonic stage depresses sonic hedgehog
RT and Hoxd-4 expression in the pharyngeal arch and induces skeletal
RT malformation in flounder (Paralicthys olivaceus) embryos.";
RL Dev. Growth Differ. 41:143-152(1999).
EMBL; AB029748; BAA82360.1; -
SQ SEQUENCE 414 AA; 45945 MW; E1FB12EE CRC32;

Query Match

Best Local Similarity 53.9%; Score 1126; DB 13; Length 414;
Matches 218; Conservative 71; Mismatches 96; Indels 10; Gaps 6;

QY 7 LRLPLCCALLALPAQSGGPGRGVRRRYARKQLVPLLYKQFVGPVPTLTGASGPAER 66
DB 9 LAGVITCLSLVS-SGKCGGPGRG-YGRRRAPKK-LTPLAYKQFIPNVAEKTLTGASGRYE 65
QY 67 VARGSERFDLVNPNNDITFKDENSGADRLMTERCKEKNALALAYNMMPGVRLRY 126
DB 66 ITRNSERFKELTPNYNDITFKDEMENTGADRLMTERCKEKNALALAYNMMPGVRLRY 125
QY 127 EGDDEGCHNAQDSLYHEGRALDITTSDDRNKYGGLARLAVEAGFDWYTESNNHAYV 186
DB 126 EGDDEGCHNEESLHYEGRAVDITTSDDRNKYGGLARLAVEAGFDWYTESNNHAYV 185
QY 187 KADNSLAVRAGGCGPGNATVRLWSGKRLRELHGRDWMVLAADASGRVPTPLFLDR 246
DB 186 KADNSLAVRAGGCGPGNATVRLWSGKRLRELHGRDWMVLAADASGRVPTPLFLDR 245
QY 247 LORRASFAVETEMPPKLLTPMHLVFAARGAP--APGDFAPFARLACGSVLAPCG 304
DB 246 STRRFLFYVLETD-SGOKITLLTAHLTFVGHSHSTERAHNGMAVYASQVRFQYTFVLD 304
QY 305 GDALRPARVAVAREEAVGAPLTAAGTLLVNDVLASCYAVLASHQMAHAFAPRLRLH 364
DB 305 AERLQGVTVARITTOHEGSEFAVTAQGVVVDVYVSCFAVYQIDHELAHMAALPVLAH 364
QY 365 ALGALL-----PGAVOPTGMHWYSLLYRLAEELL 395
DB 365 WVSLLFSSQPAQAKQDGVHWSKILYOLGTWLL 399

RESULT 3
057404

ID 057404 PRELIMINARY; PRT; 434 AA.
AC 057404; 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DE 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
GN SONIC HEDGEHOG-RELATED PROTEIN.
OS Pleurodeles waltlil (Iberian ribbed newt).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC Batrachia; Caudata; Salamandroidae; Salamandridae; Pleurodeles.
RN [1]
RP SEQUENCE FROM N.A.
RA CAUBIT X., NICOLS S., LE PARCO Y.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003532; AAB94412.1; -
DR HSSP; Q62226; 1VH.
DR PFAM; PF01085; HH_signal; 1.
DR PFAM; PF01079; Hint; 1.
DR PRINTS; PR00632; SONICHOG.
SQ SEQUENCE 434 AA; 48421 MW; AFFFD0F3 CRC32;

Query Match

Best Local Similarity 53.0%; Score 1107.5; DB 13; Length 434;
Matches 225; Conservative 59; Mismatches 91; Indels 43; Gaps 8;

QY 5 TMLPLCCALLALPAQSGGPGRGVRRRYARKQLVPLLYKQFVGPVPTLTGASGPAE 64
DB 18 TLLVPL-----GLGCGGPGRG-IGRRRPQK-LTPLAYKQFIPNVAEKTLTGASGRYE 66
QY 65 GVARASERFDLVNPNNDITFKDENSGADRLMTERCKEKNALALAYNMMPGVRLRY 124
DB 67 VITRSERFKELTPNYNDITFKDEMENTGADRLMTERCKEKNALALAYNMMPGVRLRY 126
QY 125 VEGMEDCHNAQDSLYHEGRALDITTSDDRNKYGGLARLAVEAGFDWYTESNNHAYV 184
DB 127 VEGMEDCHNEESLHYEGRAVDITTSDDRNKYGGLARLAVEAGFDWYTESNNHAYV 186
QY 185 SVKADNSLAVRAGGCGPGNATVRLWSGKRLRELHGRDWMVLAADASGRVPTPLFLD 244
DB 187 SVKADNSLAVRAGGCGPGNATVRLWSGKRLRELHGRDWMVLAADASGRVPTPLFLD 246
QY 245 ROLORRASFAVETEMPPKLLTPMHLVFAARGAPAGD-----FAPVF 290
DB 247 EETJAKFVYVLETSPLFRRLTAHLTFVAVO---EHPGNSAGNFRSKGRFRSFE 303
QY 291 ARLRAGDSVLAPG--GDALRPARVAVAREEAVGAPLTAAGTLLVNDVLASCYAVLE 348
DB 304 ASSVRGRHVLLEDEGRGLREATYDRYLEATGAYVTAHGTIVIDRYLASCYAVIE 363
QY 349 SHQMAHAFAPRLRLHAALGAL-----LPGAVOPTGMHWYSLLYRLAEELL 395
DB 364 ESHMAHAFAPRLRV--GGLTFFSFQDYSSHPAPSAQEGVHWSKILYOLGTWLL 419
RESULT 4
073803
ID 073803 PRELIMINARY; PRT; 442 AA.
AC 073803; 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DE 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
GN FUGU HEDGEHOG.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
OC Tetraodontiformes; Tetraodontidae; Tetraodontidae; Fugu.
RN [1]
RP SEQUENCE FROM N.A.
RA GELNER K., BRENNER S.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
EMBL; AF056116; AAC34384.1; -

[illegible]

RESULT	7
09WUP6	
ID	09WUP6
AC	PRELIMINARY;
DT	01-NOV-1999 (TIMBREL); 12, Created)
DT	01-NOV-1999 (TIMBREL); 12, Last sequence update)
DT	01-NOV-1999 (TIMBREL); 12, Last annotation update)
DE	DESERT HEDGEHOG PROTEIN (FRAGMENT).
CN	DHH.
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia
OC	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=SPRAGUE-DAWLEY;
RA	GARGES P.L., METER R.A. JR., BROWN C.A., PRICE D.K.:
RT	"Desert hedgehog in the rat."
RL	Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; AF148226; AACD3127.1; -
FT	NON TER
FT	NON TER
SO	SEQUENCE
	129 AA; 14578 MW; CB8B2D40 CRC32;

RESULT	8		
09MW29			
ID	09MW29	PRELIMINARY;	PRT: 177 AA.
AC	09MW29;		
DT	01-NOV-1999 (TReMBrel. 12, Created)		
DT	01-NOV-1999 (TReMBrel. 12, Last sequence update)		
DT	01-NOV-1999 (TReMBrel. 12, Last annotation update)		
DE	INDIAN HEDGEHOG PROTEIN (FRAGMENT).		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;		
OC	Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
RM	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-SPRAGUE-DAWLEY; TISSUE-KIDNEY;		
RA	GARGES P.L., MEYER R.A., JR., BROWN C.A., PRICE D.K.;		
RT	"Indian hedgehog in rat."		
RL	Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: AF162914; AAD45372.1; -.		
FT	NON_TER	1	1
FT	NON_TER	177	177
SO	SEQUENCE	177 AA;	19739 MW; 24CF1044 CRC32;

Query Match	29.9%	Score 624	DB 11	Length 177
Best Local Similarity	65.5%	Pred. No. 2	1e-41	
Matches 116	Conservative 25	Mismatches 36	Indels 0	Gaps 0
QY	18	LMTECKERVNALATAVNMMPGVRLRYTEGWDDEGHHADSLAHEGRALDITTSDDRDN	157	
DB	1	LTGQCKDRNLSTLSTVNMQMGVRLRYTEGWDDEGHHSESLAHEGRADVITTSDDRDN	60	
QY	138	KYGLLARLAVLEGFWYTYESENHYVSVKADNSLAVDAGCGFCGNATVRLMSGERKCLR	217	
DB	61	KYGLLARLAVLEGFWYTYESENKAHVCYSKSHSAAATGCGFPAGAAVHLETGERVALS	120	
QY	218	ELHRGWDVLADASGRVYPTVLTLDLSDLORBSAFVAVETEMPRKLLTPMHWLF	274	
DB	121	AKVPDRVLAMGEDGNPTFSVLLTFLDEPNRLAFQVLETODPPRLALTPAHLLF	177	

Query	Best local match	Similarity	Score	DB	Length
Db	2	YKQFIPNAEXTLGASGVEBKRTFNSRFRFLPNYPDIIFFDEENTGADRLMTORCK	61		
Qy	105	ERYNALTAIVNNMFGVRLRYTEGDEDEGHAODSLHYEGALDITTSRDRPNKYGILAR	164		
Db	62	DKLNALTAIVNDQWPGVRLRYTEGDEDEGHSSESLHYEGRAVDITTSRDRSKYGLMAR	121		
Qy	165	LAVEAGFDWYVESRNHY	182		
Db	122	LAVEAGFDWYTESKAHI	139		
RESULT	10				
Q9YGV7					
ID	Q9YGV7	PRELIMINARY:	PR:	150 AA.	
AC	Q9YGV7				
DT	01-MAY-1999	(TREMBlrel, 10, Created)			
DT	01-MAY-1999	(TREMBlrel, 10, last sequence update)			
DT	01-NOV-1999	(TREMBlrel, 12, last annotation update)			
DE	SONIC HEDGEHOG (FRAGMENT)				
OS	Bos taurus (Bovine)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos				
OC	Bovinae; Bos				
RA	SEQUENCE FROM N.A.				
RC	TISSUE=TOOTH GERM;				
RA	KOTAMA E., IWAMOTO M., OHMORI T., KURISU K., WU C., OOKURA T., BASHIR M.M., TUCKER T., PACIFICI M.;				
RT	"Development of Stratum Intermedium and its Role as a Sonic Hedgehog-Signaling structure during Odontogenesis."				
RL	Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.				
EMBL	AF144100: AAD33926.1; -				
FT	NON_TER	1	1		
FT	NON_TER	139	139		
SO	SEQUENCE	139 AA;	15961 MW;	32D3F025 CRC32;	
Query Match		29.6%;	Score 618;	DB 6;	Length 139;
Best local match		Similarity 77.5%;	Pred. NO. 4.4e-41;		
Matches 107;	Conservative	22;	Mismatches 9;	Indels 0;	Gaps 0;

RESULT	ID	Q9YGV7	PRELIMINARY:	PRT:	150 AA.
AC	Q9YGV7	01-MAY-1999	(TRIMBLrel. 10, Created)		
DT	01-MAY-1999	(TRIMBLrel. 10, Last sequence update)			
DT	01-NOV-1999	(TRIMBLrel. 12, Last annotation update)			
DE	SONIC HEADSHOG (FRAGMENT)				
OS	Ambystoma mexicanum (Axolotl)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;				
CC	Batrachia; Caudata; Salamandroidea; Ambystomatidae; Ambystoma.				
RN	[1]				

RP SEQUENCE FROM N.A.
RA TOBOK M.A., IZPITZUA-BELMONTE J.C., GARDINER D.M., BRYANT S.V.;
RL submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
R1 EMBL; AF031480; AAD18128.1; -
DR HSSP; Q62226; 1VHH.
FT NON TER 1
FT 150 1
FT NON TER 150
SO SEQUENCE 150 AA; 16599 MW; 9356329B CRC32;

Query Match	29.0%	Score 606;	DB 13;	length 150;
Best Local Similarity	72.0%	Pred. No. 4.2e+40;		
Matches 108;	Conservative 25;	Mismatches 17;	Indels 0;	Gaps 0;

```

0Y 84 DIFDEENSAGDRMLTMRCECRVVALAIYANNMPGRLVTEGMDGHDHADSHTYE 143
Db 1 DIFDEENTGADRLMTROCRDKLNTALAIYVNMOPGKLVTEGMDGHDHSESLHTYE 60
0Y 144 GRADITSDSDRNNKYGTLALAYEAGDWMYYSRRNNHVSYKADSLAYRAGGCEPFGN 203
Db 61 GRADITISDRKRSTYGTMLARLAYEAGDWMYFESKALHICVSYAKESVNAKSGCCPPAS 120

```

QY 204 ATVRLMSGERRKGLRELHRGDVLAADASGR 233
| | | : : : | | | | | |
Db 121 AKYTLHGVTSPVKDLRPGDRVLAADGQGR 150

RESULT	11	
09W6C1		
ID	09W6C1	PRELIMINARY;
AC	09W6C1.	PRT; 138 AA

DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
DE SONIC HEDGEHOG PROTEIN (FRAGMENT).
Cntr

05. *Eleutherodactylus coqui*.
06. Eukaryota; Metazoa; Chordata; Craniata; Amphibia.
0C. Batrachia; Anura; Neobatrachia; Bufonidae; Leptodactylidae;
0C. Eleutherodactylus.

RA CARL T.F., RICHARDSON M.K., OLSSON L., SCHLOSSER G., KLYMKOWSKY M.W.,
RA HANKEN J.;
RT "Differences in vertebrate limb development revealed by studies of the
RT direct developing frog *E. coqui*.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DE FMR1, A0113403, A0273436, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 8

FI	NON_ITER	1	1
FT	NON_ITER	138	138
SQ	SEQUENCE	138 AA;	15751 MW; 2D3E8060 CRC32

Query Match	28.7%;	Score 599;	DB 13;	Length 138;
Best Local Similarity	75.4%;	Pred. No. 1.3e-39;		
Matches 104;	Conservative 23;	Mismatches 11;	Indels	

49 VPGVVERTIGASGPAEGRVARSSERFRDLYPNYNPDITFKDEENSGADRLMTERCKERVN 108
QY
1 IPNVAERTIGASGVEGKITRNSERREKELTPNYSNDITFKDEENTADRLMTOPECKDKIN 60
Db

[illegible]

```

Db 121 AGFDWVYESKAHICSV 138

```

RESULT	12
096699	
ID	096699
PRELIMINARY;	
PRT;	185 AA

AC 096699;
DT 01-MAY-1999 (TrEMBLrel, 10, Created)
DT 01-MAY-1999 (TrEMBLrel, 10, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel, 12, Last annotation update)
DE HEDGEHOG PROTEIN (FRAGMENT).

05 *Junonia coenia* (Peacock butterfly) (Pieris coenia).
0C Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
0C Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia.
0C Papilionoidea; Nymphalidae; Nymphalinae; Junonia.

RA SEQUENCE FROM N.A.
RP KEYS D.N., LEWIS D.L., SELEGUE J.E., PEARSON B.J., GOODRICH L.V.,
RA JOHNSON R.L., GATES J., SCOTT M.P., CARROLL S.B.;
RT "Recruitment of a hedgehog regulatory circuit in butterfly eyespot
evolution".

```

RL Science 0:0-0(1999).
DR EMBL; AF117742: MAD08931.1; -.
DR HSSP; 062226; 1VHH.
FT
FT     NON_TER      1
FT     NON_TER      185
SQ SEQUENCE      185 AA; 20745 MW; 96A09B5A CRC32;

```

Query Match	28.0%	Score 585	DB 5	Length 185
Best Local Similarity	64.1%	Pred. No. 2.4e-38		
Matches 109	Conservative 26	Mismatches 33	Indels 2	Gaps 2

```

Oy  59 ASGPAGEVRVANGSEREPDLPVNPNDIIFKDEENSGADRINTRECKERVNALAIAYNNMM 118
    ||| ||| : :||| ||| ||| : ||| ||| ||| : ||| |||
Db   3 ASGPPEGRIITDDEKFRDLPVNPNDIDIFKDEGTGADRIMTQRCHEKLNLTALISVNNQW 62

```

Oy 119 PGVRLVTEGDEWDGHHAQDSLHYEGRALDITTESDRDRNKGGLLARLAVENGDFWYYES 178
| | | | : | : | | | | : | | | | | : | | :
Db 63 PGVRLVTEGDEWDENSHDNSLHYEGRAVDLTTSRDRSHKNGMLARLAVENGDFWVEYEN 122

QY 179 RNHVHVSVKADNSLAVRAGCFEPGNATVRLMSGERKGRELHRGDWTLAA 228
::: | ||| ::| : | : | :
Db 123 RSYIHCSVKTESSTGTGA-GCFFSGCAVHTENGP-XDIALSLKGNKVLA 1700

RESULT	13	
042128		
ID	042128	PRELIMINARY;
10	042128	PRT; 119 AA.

DT 01-JAN-1998 (TREMblel. 05, Created)
DT 01-JAN-1998 (TREMblel. 05, Last sequence update)
DT 01-NOV-1999 (TREMblel. 12, Last annotation update)
DE SHH, PARTIAL CDS (FRAGMENT).

0C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
0C Neopterygia; Teleostei; Euteleostei; Acanthopterygii; Atherinomorpha;
0C Cyprinodontiformes; Adrianichthyidae; Adrianichthyidae; Orziniinae;
0C Orzias.

RP SEQUENCE FROM N.A.
RC STRAIN=BBRR; TISSUE=WHOLE EMBRYO;
RA ARAKI K.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases
EMBL accession: AB022661.1

FT	NON_TER	1	1
FT	NON_TER	119	119
SO	SEQUENCE	119 AA;	13179 MM; 1CAE5021 CRC32

Query Match	20.68	Score 431	DB 13	Length 119
Best Local Similarity	64.78	Pred. No. 1.2e-26		
Matches 77; Conservative	23	Mismatches 19	Indels 0	Gaps 0

QY	124	RVTEGDEDEGHNAODSLHYEGRALDITTSBDRNRKYGTLARLAVEAGFDWVYYEERNHHV	183

THIS PAGE BLANK (uspto)

THIS PAGE BLANK

Db	61	GPABGRVYRSGSERRRDLVPPNPNPILIFKRDENSQADMLMTERCKERNALALIVAMNMWPG	120
QY	121	VLRLVTEGMDGEDGHHADDSLHYEGSRALDITTSDDRNNKYGGLARLAVEAGFDWYYESRN	180
Db	121	VLRLVTEGMDGEDGHHADDSLHYEGSRALDITTSDDRNNKYGGLARLAVEAGFDWYYESRN	180
QY	181	HHVHSVKADNSLAVRAGCGFCGNAITVRLMSGKGLRELRHGMVYLAADASGRVPTPVY	240
Db	181	HHVHSVKADNSLAVRAGCGFCGNAITVRLMSGKGLRELRHGMVYLAADASGRVPTPVY	240
QY	241	LFELDRDLORRASFAVETEMPPRKLLLTPMHLVFAAGSPAPAGDFEAPVARRLLRAGDSY	300
Db	241	LFELDRDLORRASFAVETEMPPRKLLLTPMHLVFAAGSPAPAGDFEAPVARRLLRAGDSY	300
QY	301	LAPGCDALRPARVARVARAAEAVGVFAPLTAHGTLLVNDVTLASCYAVLESHQMAHRAAPL	360
Db	301	LAPGCDALRPARVARVARAAEAVGVFAPLTAHGTLLVNDVTLASCYAVLESHQMAHRAAPL	360
QY	361	RLHHLALGALLPGAGVQPTGMHWYSRLLYRLAEELLG	396
Db	361	RLHHLALGALLPGAGVQPTGMHWYSRLLYRLAEELLG	396

```

1      RESULT      2
2      US-08-356-060A-9
3      Sequence 9, Application US/08356060A
4      Patent No. 5844079
5      GENERAL INFORMATION:
6      APPLICANT: Ingham, Phillip W.
7      APPLICANT: McMahon, Andrew P.
8      APPLICANT: Tablin, Clifford J.
9      TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
10     TITLE OF INVENTION: Proteins and Uses Related Thereto
11     NUMBER OF SEQUENCES: 47
12     CORRESPONDENCE ADDRESSES:
13     ADDRESSEE: LAHIVE & COCKFIELD
14     STREET: 60 State Street
15     CITY: Boston
16     STATE: MA
17     COUNTRY: USA
18     ZIP: 02109
19     COMPUTER READABLE FORM:
20     MEDIUM TYPE: Floppy disk
21     COMPUTER: IBM PC Compatible
22     OPERATING SYSTEM: PC-DOS/MS-DOS
23     SOFTWARE: ASCII(text)
24     CURRENT APPLICATION DATA:
25     APPLICATION NUMBER: US/08/356,060A
26     FILING DATE: 14-DEC-1994
27     CLASSIFICATION: 435
28     PRIOR APPLICATION DATA:
29     APPLICATION NUMBER: US 08/176,427
30     FILING DATE: 30-DEC-1993
31     ATTORNEY/AGENT INFORMATION:
32     NAME: Vincent, Matthew P.
33     REGISTRATION NUMBER: 36,709
34     REFERENCE/DOCKET NUMBER: HMT-006CP
35     TELECOMMUNICATION INFORMATION:
36     TELEPHONE: (617) 227-7400
37     TELEFAX: (617) 227-5941
38     INFORMATION FOR SEQ ID NO: 9:
39     SEQUENCE CHARACTERISTICS:
40     LENGTH: 396 amino acids
41     TYPE: amino acid
42     TOPOLOGY: linear
43     MOLECULE TYPE: protein
44     US-08-356-060A-9

```

Query Match	96.2%;	Score 208;	DB 2;	Length 396;
Best Local Similarity	96.5%;	Pred. No. 2.5e-214;		
Matches 382;	Conservative	6;	Mismatches 8;	Indels 0;
			Gaps	0

QY	1	MAATNLLPLOCCLALLLPLPOSGEGGPGVGRRRYARKOKPLLYLKQFVGVBERLTGAS	60
		:	
Db	1	MALPASLLPLOCCLALLLPLSASCGPGGPGVGRRRYARKOKPLLYLKQFVSPMBRTLGAS	60
QY	61	GPAEGRAVARGSEFRDLYPNVNPDIIFKDEBNSGADRLMTERCKERVNALAIWNNMPPG	120
Db	61	GPAEGRAVTRRSEFRRLDLYPNVNPDIIFKDEBNSGADRLMTERCKERVNALAIWNNMPPG	120
QY	121	VLRLVTEGWDGEGHHAODLSLHYESRALDITTSDDRNNKYGLLARLAVEAGFDWYYESRN	180
Db	121	VLRLVTEGWDGEGHHAODLSLHYESRALDITTSDDRNNKYGLLARLAVEAGFDWYYESRN	180
QY	181	HHVHSVADNSLAVRACGCFPGNATVYLSMGERGKLELIRGDMVYLAADASGRVPPVYL	240
Db	181	HHVHSVADNSLAVRACGCFPGNATVYLSRGERGKLELIRGDMVYLAADASGRVPPVYL	240
QY	241	LEFLRDLORRASFEVAVETEMPPRKLTLTPHNLVFAARGPAPAPGDFEAPVEFARRLACDSV	300
Db	241	LEFLRDLORRASFEVAVETEMPPRKLTLTPHNLVFAARGPAPAPGDFEAPVARRRLACDSV	300
QY	301	LAPGCDALPRPARVARVAREEAVGFAPLTHGTLVLNDVTLASCAYLVESHQMAHRAAPL	360
Db	301	LAPGCDALOPARVARVAREEAVGFAPLTHGTLVLNDVTLASCAYLVESHQMAHRAAPL	360
QY	361	RLTHALGALLPGSAVOPTGMHWTSRLLYRLAEELTG	396
Db	361	RLTHALGALLPGSAVOPTGMHWTSRLLYRLAEELTG	396

```

RESULT 3
US-08-176-427B-8
Sequence 8, Application US/08176427B
Patent No. 5789543
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tabin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
TITLE OF INVENTION: Proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,427B
FILING DATE: 30-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMI-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-176-427B-8

```

MOLECULE TYPE: protein
US-08-176-427B-8

Query Match 58.3%; Score 1217; DB 1; Length 437;
 Best Local Similarity 58.7%; Pred. No. 1.6e-126;
 Matches 249; Conservative 53; Mismatches 82; Indels 40; Gaps 12;

```

QY 7 LRLPCCALTA-----LPASCGRGPGVGRRRARQVLYLVKQVPGVPERLTASG 61
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4 LLARCFVLIASSLVCPGLACGPGRG-FGKRHPKR-LTPLAVKQFIIPNVAERTLTASG 61
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 62 PAEGRVARGSERFDLVNPNPDIIFFDEENSGADRLTERCKERVNALAIVANNMPPGV 121
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 62 RYEGKITRNSRFKELPNPNPDIIFFDEENTGADRLMTQCKKLNALASVNNQMPGV 121
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 122 RLRTBEGWDEGHHADSLHIEGRALDITTSDRKNKYGILLARLAVAGFDWYYESRNH 181
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 122 RLRTBEGWDEGHHESLSHIEGRAVDITTSDRSKYGMARLAVAGFDWYYESRAH 181
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 182 VHSVKAADNSLAVAGCGFPGNATVRLMSGKRLRELRHGDWVLADASGRVPTPVLL 241
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 182 IHCSVKAENSVAAGSGCFPGSATVHLEOGSTKLVKDLRPDRVLAADQGRLLYSDFLT 241
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 242 FLDRDLORASFAVETEMPRLKLLTPWMLVFNA---RGAPAPGDFAFVFAARRLRAG 297
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 242 FLDRDEGAKKFYIETLEPRERLLTAHLLFVAPHNDSGPTGP---SALFASRVAPG 298
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 298 DSVLA---PGGD-ALRAVARVA-REBAGVFAPLTAHGLLVNDVLAACYAVLESQHW 352
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 299 QRVYVAERGGDRLLPAVAHVTLRBEAGAYAPLTAHGLILINRYLASCYAVIEESHW 358
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 353 AHRFAFAPRLHA-LGALLP-----GGAV-----OPT-GMHYSRLLYLA 391
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 359 AHRFAFAPRLHALLALAPARTDGGGGSIPAAOSATEARGABPTAGIHYSOLLYHIG 418
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 392 EELL 395
    ||
Db 419 TWLL 422
  
```

RESULT 4

US-08-356-060A-11
 ; Sequence 11, Application US/08356060A
 ; Patent No. 5844079
 ; GENERAL INFORMATION:
 ; APPLICANT: Ingham, Phillip W.
 ; APPLICANT: McMahon, Andrew P.
 ; APPLICANT: Tablin, Clifford J.
 ; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
 ; NUMBER OF SEQUENCES: 47
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 60 State Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: ASCII(text)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/356,060A
 ; FILING DATE: 14-DEC-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/176,427
 ; FILING DATE: 30-DEC-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Vincent, Matthew P.
 ; REGISTRATION NUMBER: 36,709
 ; REFERENCE/DOCKET NUMBER: HMI-006CP
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941
 ; INFORMATION FOR SEQ ID NO: 11:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 437 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-356-060A-11.

Query Match 58.3%; Score 1217; DB 2; Length 437;
 Best Local Similarity 58.7%; Pred. No. 1.6e-126;
 Matches 249; Conservative 53; Mismatches 82; Indels 40; Gaps 12;

```

QY 7 LRLPCCALTA-----LPASCGRGPGVGRRRARQVLYLVKQVPGVPERLTASG 61
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4 LLARCFVLIASSLVCPGLACGPGRG-FGKRHPKR-LTPLAVKQFIIPNVAERTLTASG 61
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 62 PAEGRVARGSERFDLVNPNPDIIFFDEENSGADRLTERCKERVNALAIVANNMPPGV 121
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 62 RYEGKITRNSRFKELPNPNPDIIFFDEENTGADRLMTQCKKLNALASVNNQMPGV 121
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 122 RLRTBEGWDEGHHADSLHIEGRALDITTSDRKNKYGILLARLAVAGFDWYYESRNH 181
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 122 RLRTBEGWDEGHHESLSHIEGRAVDITTSDRSKYGMARLAVAGFDWYYESRAH 181
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 182 VHSVKAADNSLAVAGCGFPGNATVRLMSGKRLRELRHGDWVLADASGRVPTPVLL 241
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 182 IHCSVKAENSVAAGSGCFPGSATVHLEOGSTKLVKDLRPDRVLAADQGRLLYSDFLT 241
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 242 FLDRDLORASFAVETEMPRLKLLTPWMLVFNA---RGAPAPGDFAFVFAARRLRAG 297
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 242 FLDRDEGAKKFYIETLEPRERLLTAHLLFVAPHNDSGPTGP---SALFASRVAPG 298
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 298 DSVLA---PGGD-ALRAVARVA-REBAGVFAPLTAHGLLVNDVLAACYAVLESQHW 352
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 299 QRVYVAERGGDRLLPAVAHVTLRBEAGAYAPLTAHGLILINRYLASCYAVIEESHW 358
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 353 AHRFAFAPRLHA-LGALLP-----GGAV-----OPT-GMHYSRLLYLA 391
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 359 AHRFAFAPRLHALLALAPARTDGGGGSIPAAOSATEARGABPTAGIHYSOLLYHIG 418
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 392 EELL 395
    ||
Db 419 TWLL 422
  
```

RESULT 5

PCT-US95-15463-20
 ; Sequence 20, Application PC/TUS9515463
 ; GENERAL INFORMATION:
 ; APPLICANT: The Johns Hopkins University School of Medicine
 ; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 4225 Executive Square, Suite 1400
 ; CITY: La Jolla
 ; STATE: CA
 ; COUNTRY: U.S.A.
 ; ZIP: 92037
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/15463
 ; FILING DATE: 01-DEC-1995
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Haile, Lisa A.
 ; REGISTRATION NUMBER: 38,347

REFERENCE/DOCKET NUMBER: 07265/080W01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-15463-20

Query Match 58.1%; Score 1214; DB 4; Length 437;
Best Local Similarity 58.5%; Pred. No. 3.5e-126;
Matches 248; Conservative 54; Mismatches 82; Indels 40; Gaps 12;

QY 7 LPLCLALLA-----LPAQSCGPGRGVGRRRYARKOLVPLLYKQFVGVPERTLGASG 61
DB 4 LLARCELYVLAASLLVCPGLACGPRG-FGKRHRPK-LTPLAYKQFIPVNAEKTIGASG 61
QY 62 PAEGVARGSERFDLVPNYNPDIIIFKDEENSGADRLMTERCKERYNALAIANNMPGV 121
DB 62 RYGGKTRNSERKELTPYNNPDIIIFKDEENSGADRLMTERCKERYNALAIANNMPGV 121
QY 122 RLRTGEMWEDGHHAQDSLYEGRALDITSDRDNKYGILLARLAVEGFDWVYESRNH 181
DB 122 KLRTGEMWEDGHHEESLHTEGRAVDITSDRDNKYGILLARLAVEGFDWVYESRNH 181
QY 182 VHVSKADNSLAVRAGCGPGNATVRLMSGERKGLRELHNGDMVLAADASGRVPTPYLL 241
DB 182 IHCVSVAENSVAAGSGCGPGSATVHLEGGTKLVKDLRPGDVLAAADQGRLLYSDFLT 241
QY 242 FLDRDLQRRASFVAVETEMPRKLLTPMHLVFAA---RGPAAPGDFAPVAFARLRAG 297
DB 242 FLDRDGAQKRVVETLEPRERLLTAHLFLVAPHNDGPTPGP---SALFASRVRG 298
QY 298 DSVLA---PGCD-ALRPARYAVA-REAVGVAPLTAAGTLVNDVLAACVAVLESHQW 352
DB 299 QRYVVAERGGDRRLPAVHVSVTLREEGAVAPLTAAGTLINRVLAACVAVLESHQW 358
QY 353 AHRAFAFLRLHA-LGALLP-----GGAV-----QPT-GMHWYRLLYRLA 391
DB 353 AHRAFAFLRLHA-LGALLP-----GGAV-----QPT-GMHWYRLLYRLA 391
QY 392 EELL 395
DB 419 TWLL 422

RESULT 6
PCT-US95-15923-20
Sequence 20, Application PC/TUS9515923
GENERAL INFORMATION:
APPLICANT: The Johns Hopkins University School of Medicine, et al.
TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: LA JOLLA
STATE: CA
COUNTRY: U.S.A.
ZIP: 92037
MEDIUM TYPE: floppy disk
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/15923
FILING DATE: 04-DEC-1995
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/043W01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-15923-20

Query Match 58.1%; Score 1214; DB 4; Length 437;
Best Local Similarity 58.5%; Pred. No. 3.5e-126;
Matches 248; Conservative 54; Mismatches 82; Indels 40; Gaps 12;

QY 7 LPLCLALLA-----LPAQSCGPGRGVGRRRYARKOLVPLLYKQFVGVPERTLGASG 61
DB 4 LLARCELYVLAASLLVCPGLACGPRG-FGKRHRPK-LTPLAYKQFIPVNAEKTIGASG 61
QY 62 PAEGVARGSERFDLVPNYNPDIIIFKDEENSGADRLMTERCKERYNALAIANNMPGV 121
DB 62 RYGGKTRNSERKELTPYNNPDIIIFKDEENSGADRLMTERCKERYNALAIANNMPGV 121
QY 122 RLRTGEMWEDGHHAQDSLYEGRALDITSDRDNKYGILLARLAVEGFDWVYESRNH 181
DB 122 KLRTGEMWEDGHHEESLHTEGRAVDITSDRDNKYGILLARLAVEGFDWVYESRNH 181
QY 182 VHVSKADNSLAVRAGCGPGNATVRLMSGERKGLRELHNGDMVLAADASGRVPTPYLL 241
DB 182 IHCVSVAENSVAAGSGCGPGSATVHLEGGTKLVKDLRPGDVLAAADQGRLLYSDFLT 241
QY 242 FLDRDLQRRASFVAVETEMPRKLLTPMHLVFAA---RGPAAPGDFAPVAFARLRAG 297
DB 242 FLDRDGAQKRVVETLEPRERLLTAHLFLVAPHNDGPTPGP---SALFASRVRG 298
QY 298 DSVLA---PGCD-ALRPARYAVA-REAVGVAPLTAAGTLVNDVLAACVAVLESHQW 352
DB 299 QRYVVAERGGDRRLPAVHVSVTLREEGAVAPLTAAGTLINRVLAACVAVLESHQW 358
QY 353 AHRAFAFLRLHA-LGALLP-----GGAV-----QPT-GMHWYRLLYRLA 391
DB 353 AHRAFAFLRLHA-LGALLP-----GGAV-----QPT-GMHWYRLLYRLA 391
QY 392 EELL 395
DB 419 TWLL 422

RESULT 7
PCT-US95-02315-2
Sequence 2, Application PC/TUS9502315
GENERAL INFORMATION:
APPLICANT: Jessell, Thomas M.
APPLICANT: Dodd, Jane
APPLICANT: Roelink, Henk
APPLICANT: Edlund, Thomas
TITLE OF INVENTION: DNA ENCODING A VERTEBRATE HOMOLOG OF
TITLE OF INVENTION: HEDGEHOG, VHH-1, EXPRESSED BY THE NOTOCHORD, AND USES
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02315
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: John P. White
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 45375-A-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-02315-2

```

Query Match	57.68%	Score 1203;	DB 4;	Length 437;
Best Local Similarity	57.88%	Pred. No. 5.7e-125;		
Matches 245;	Conservative 52;	Mismatches 87;	Indels 40;	Gaps 10

QY	7	LLPCLC-----ALLAPAOCSGGRGVGRRRARAQOLVLTQFVPGVPTLGLASG	61
Db	4	LLANGELVALLASSLLVCGCLAGCPGRG-FGRKHQPK-LTPLAQKQIPINVAEXTLGLASG	61
QY	62	PAEGRVAGSERFDLPVNPNDIIFKDEENSGADRLTERCKERVALAIAVNMMPGV	121
Db	62	RYEKRIETNSREFELTPNPNDIIFKDEENTGADRLMTQCKDKMLAISVNMOPGV	121
QY	122	RLRTEGMDDEGHHAAOSLHYEGRALDITTSDBDRNRYGILRLAVAGADWYEEGRNH	181
Db	122	KLRTGEGMDGHHSEELTEGRAVDITTSDBDRSRXGMLARLAVEGDPMWYEEKAR	181
QY	182	VHVSVKADNSLAVAGGCFGNATVRLMSGERKGLREIHRDWDVLADASGRVVPYVL	241
Db	182	IHCYSKAKENSVAAKSDGCFGSATVHLNEGQGTKLVKDLSPEDRYLADDDGRLLYSDFLT	241
QY	242	FLDRDLQRRASFVAETEMPPRKLILTPMHLVFA---KGPAPGDPAPVFAARLRAG	297
Db	242	FLDRDEGAKKVFYVIEETREPERERILLTAHILLFPAAPHNDSGPTPGP---SPLFSRYRPG	298
QY	298	DSVLA---PGGD-ALRPAPARVA-REEAVGVAPLTAHGTLVNDVLASCYAVLESQHW	352
Db	299	QRYVYVAERGGDRLLPRAVHVSYTLREEBAGATYAPLTDGTLILNRYLASCITAVIEEHSW	353
QY	353	AHRAFAPLRLHLA-LGALLPG-----GAVOPTGMHWYSLRLYRLA	391
Db	359	AHRAFAFAPRLAHALLAALAPARTDGGGGSIPAPQSVAEARGAGPAPGIMHYSQLYHIG	418
QY	392	DELL 395	
Db	419	TWLL 422	

RESULT 8
 US-08-176-427B-2
 : Sequence 2, Application US/08176427B
 : Patent No. 5789543
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Ingham, Phillip W.
 : APPLICANT: McMahon, Andrew P.
 : APPLICANT: Tabin, Clifford J.
 : TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
 : TITLE OF INVENTION: Proteins and Uses Related Thereto
 :
 : NUMBER OF SEQUENCES: 33
 :
 : CORRESPONDENCE ADDRESS:
 :

```

ADDRESS: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,427B
FILING DATE: 30-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HWI-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 425 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-176-427B-2

```

Query Match	57.4%	Score 1198:	DB 1:	Length 425,
Best Local Similarity	59.4%	Pred. No. 2e-124:		
Matches 244:	Conservative 56:	Mismatches 91:	Indels 20:	Gaps 10

```

OY 1 MALLNLLP---LCCLLALLPAGSCGPGPGRARRAKOVPLLYLKFQVGVPERLT 57
    | | | | | - - - | | | : | | | | | : | | | | | : | | | | | : | |
Db 4 MLLTLRILVGFIC--ALLVSSGLTCGPBGK-IGKRHPKK-LTPLAYKQFIYNAVKTL 59
OY 58 GASPAEGRVARGSERFERDLPVNPNDIIFKDEENSGADRLMTERCKERYNALALAVMM 117
    | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db 60 GASGVEEKKITRNERRERKELTPVNPNDIIFKDEENTGADRLMQRCKDKLNALAISVMQ 119
OY 118 WPGVRLKRYTECWDGDGHHAODSLAYEBERALDITTSDBDRNKYGLLARLAYAGEFDWYTE 177
    | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db 120 WPGVKLRYTECWDGDGHHSSESLYEBERAVDITTSDBDRNKYGLMARLAYAGEFDWYTE 179
OY 178 SRNHVHSVKAINDLAYRAGGCEFGNATVYLMSEGRGLRELRGMDVULAADSGRVPT 237
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 180 SKATHCSVKAKENSVAAKSGGCEFGSATHHEHCGITLVLDLSBQDVLAAADGRLIYS 239
OY 238 PVLLEFLDRDLORASFYAVELTEMPRLKLLTPMHLVPA--RGAPAPAGDFA--PVFARR 293
    | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db 240 DELFLFDBMDSRKLFIYIEHROPRARLLLTALILRLVVAQHNQSEATGSTSQALFASN 299
OY 294 LRAGDS--VLAPGDALRPARVAVNA--REBAVGYEARLTHGTLVNDVYASCTAVLESH 350
    : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 300 VKPGRVYVLDEGGQOQLLPASVHSVSLREBASGAYABLTAGGTLINRVLASCAVIEEH 359
OY 351 OMAHAPAPRLTLHA-LGALLPGAV-----QPTGMVYSRLYRLADELL 395
    | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db 360 SMAHNAEPFRLLAAGLIALCDBDALPAAITTTGTHIMYSFLYRISQWV 410

```

RESULT 9
 US-08-356-060A-8
 Sequence 8, Application US/08356060A
 Patent No. 5844079
 GENERAL INFORMATION:
 APPLICANT: Ingham, Phillip W.
 APPLICANT: McMahon, Andrew P.
 APPLICANT: Tabin, Clifford J.
 TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing

Db 418 PGADAPGAGATAGIHWYSQLYQIGTWL 447

RESULT 11
US-08-748-591-9

; Sequence 9, Application US/08748591
; Patent No. 5759811
; GENERAL INFORMATION:
; APPLICANT: Epstein, Ervin
; APPLICANT: Hu, Zhilan
; APPLICANT: Bonifas, Jeanette
; TITLE OF INVENTION: Mutant Human Hedgehog Gene
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish and Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/748,591
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: 06510/067001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 322-5070
; TELEFAX: (415) 854-0875
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-748-591-9

Query Match 56.4%; Score 1177; DB 1; Length 462;
Best Local Similarity 54.4%; Pred. No. 4.8e-122;
Matches 245; Conservative 54; Mismatches 93; Indels 58; Gaps 10;

Db 1 MALLTNLPICCL-ALLALPAQSCGPGRGVGRRRYARKOLVPLLYKQFVGPVPEPTLGA 59
1 MLILARCLLVVSSLVCSGLACGPGRG-FGKRHRPK-LTPLAYKQFIPNVAEKTLLGA 58
QY 60 SGAPEGVANGSERFRLVNYNDIIFKDENSGADRLMTERCKERNALALAVNMMP 119
119 GVKLRVTEGDEGHSHSEELHYEGRAVDITTSRDRSKYGMALARLAVEAGFDWVYYESK 178
QY 120 GVKLRVTEGDEGHSHSEELHYEGRAVDITTSRDRSKYGMALARLAVEAGFDWVYYESK 179
119 GVKLRVTEGDEGHSHSEELHYEGRAVDITTSRDRSKYGMALARLAVEAGFDWVYYESK 178
QY 180 NHVAVSKADNSLAVRAGGCPGNATVRLMSEGRGLRELRGDMVYLAADSGHVPTPV 239
179 AHICSTKAENSVAAGSGGCPGSAIVHLDEGGTKLVKDLSPGRVLAADQGRLLSDF 238
QY 240 LFLDLRDLORRASEFAVETMPKRLLTTPWHLVFAA-----RGPAPAG 284
239 LFLDLRDLORRASEFAVETMPKRLLTTPWHLVFAA-----RGPAPAG 284
QY 285 DFAP--VFARLRAGDSVLA-----PGDALPAPAVARYA-REEVGVFAPLTAHGTLLVN 337

Db 298 ALGPRLAFASRVPRGQRYVVAERDGRLLPAAVHSVTISEAAGAYADLTAGTLLIN 357
QY 338 DYLAACVAVLESHQWARRAFAPRLTLHALGALL----- 371
Db 358 RVLASCAVIEEHSMAHRAFPRLAHALLAALAPARTDGRGSGGGGGRGGRVALTA 417
QY 371 PGAVOP-----TGMWYSRLYRLAEELL 395
Db 418 PGADAPGAGATAGIHWYSQLYQIGTWL 447

RESULT 12

US-08-356-060A-13
; Sequence 13, Application US/08356060A
; Patent No. 5844079
; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tabin, Clifford J.
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,060A
; FILING DATE: 14-DEC-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,427
; FILING DATE: 30-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-006CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-356-060A-13

Query Match 56.4%; Score 1177; DB 2; Length 475;
Best Local Similarity 54.4%; Pred. No. 5e-122;
Matches 245; Conservative 54; Mismatches 93; Indels 58; Gaps 10;

Db 1 MALLTNLPICCL-ALLALPAQSCGPGRGVGRRRYARKOLVPLLYKQFVGPVPEPTLGA 59
1 MLILARCLLVVSSLVCSGLACGPGRG-FGKRHRPK-LTPLAYKQFIPNVAEKTLLGA 58
QY 60 SGAPEGVANGSERFRLVNYNDIIFKDENSGADRLMTERCKERNALALAVNMMP 119
119 GVKLRVTEGDEGHSHSEELHYEGRAVDITTSRDRSKYGMALARLAVEAGFDWVYYESK 178
QY 120 GVKLRVTEGDEGHSHSEELHYEGRAVDITTSRDRSKYGMALARLAVEAGFDWVYYESK 179
119 GVKLRVTEGDEGHSHSEELHYEGRAVDITTSRDRSKYGMALARLAVEAGFDWVYYESK 178

```
QY      118  MPGRVLRTEEMDEMDGHHNAOOSLYIEBGRALDITTSDDSDRNKYGILATLAEAGCDWYYE   177
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      120  WPGVMRLRTEEMDEMDGHHKSLESLEYEGRANDITTSDDRKRSTGYMLAKLAEAGCDWYYE   179
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      178  SRNHVSVSKADNSLAVRAGCGCFPGNATVRLMSCGERKKRELIHMGDWYLALDASGRVPT    237
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      180  SKAHI -CSVKAEHSVAASKSGCGSPGASTVHLHEHGTKVLKDLSHGDVYLAAADAGRLLVS   236
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      238  PVLV-ELDRDLDQRASAFNAVETEMPBRKLTLTPHYLFAN--RRPARAPQDFA--PVFAR    292
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      239  DFLFLFDLRMDSRKLEFYIETTRPRRLLTAAHLIFVAPQHQSATSGTSSQALFAS    296
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      293  RLRAQGS--VLAPRGDALRPARAARVA-RREEAVGFAPLAHTGLLVNDVASCYAVALS    348
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      299  NVKRGQPVVVLVEGGQQQLPPLASVHSVSLREBASCAVAPTACGFIILNRVLASCAYIVEE    350
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      350  HQMAHRAPAPRLRLHA-LGALLPGGAV-----OPTGMHWTSRLRYRLABELL    395
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      359  HSMHMAAAPPRRLAQGLLAACPBGALPTATTGTHWSRSRLRYRGISGWL    410
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 14
 PCT-US95-15923-19
 Sequence 19, Application PC/TUS9515923
 GENERAL INFORMATION:
 APPLICANT: The Johns Hopkins University School of Medicine, et al.
 TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 4225 Executive Square, Suite 1400
 CITY: La Jolla
 STATE: CA
 COUNTRY: U.S.A.
 ZIP: 92037
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/15923
 FILING DATE: 04-DEC-1995
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Haile, Lisa A.
 REGISTRATION NUMBER: 38,347
 REFERENCE/DOCKET NUMBER: 07265/043WO1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 619/678-5070
 TELEFAX: 619/678-5099
 INFORMATION FOR SEQ ID NO: 19:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 425 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: both
 MOLECULE TYPE: protein
 PCT-US95-15923-19

Query Match	55.5%	Score 1158:	DB 4:	Length 425:
Best Local Similarity	59.0%:	Pred. No. 5.4e-120:		
Matches 243:	Conservative 54:	Mismatches 92:	Indels 22:	Gaps 12:
QY	1	MALTLNLP----	LCCLALLALPAQSCGGRGVRGRRRARAKOVLVLTIKQFIPGVPERTL	57
Db	4	MLLRLRLVGFIC--	ALLVSSGLTCGGGGRG-IGRRRHPKK-LTFLAVKQFIPNVAEKL	59
QY	58	GASGAEGRVARGSRFEDVLPVNPDIILFDEENSGADRLMTECKEKENVALATAVMM	117	
	60	GASGYEETIRNSRFEKLIPVNPDIILFDEENTGADRLMTECKEKKALATAVNMC	119	

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2000, 07:29:48 ; Search time 1833.88 Seconds

(without alignments)
-631.772 Million cell updates/sec

Title: US-08-900-220-8

Perfect score: 1191
Sequence: 1 ATGCTCTCTGACCAATCT.....CGAGAGACTACTGGCTGA 1191

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 1765538

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl:*

1: gb_da1:*
2: gb_da2:*
3: gb_om:*
4: gb_om:*
5: gb_pat:*
6: gb_ph:*
7: gb_ph:*
8: gb_ph:*
9: gb_ph:*
10: gb_ph:*
11: gb_ph:*
12: gb_ph:*
13: gb_ph:*
14: gb_ph:*
15: gb_ph:*
16: gb_ph:*
17: gb_ph:*
18: gb_ph:*
19: gb_ph:*
20: gb_ph:*
21: gb_ph:*
22: gb_ph:*
23: gb_ph:*
24: gb_ph:*
25: gb_ph:*
26: gb_ph:*
27: gb_ph:*
28: gb_ph:*
29: gb_ph:*
30: gb_ph:*
31: gb_ph:*
32: gb_ph:*
33: gb_ph:*
34: gb_ph:*
35: gb_ph:*
36: gb_ph:*
37: gb_ph:*
38: gb_ph:*
39: gb_ph:*
40: gb_ph:*
41: gb_ph:*
42: gb_ph:*
43: gb_ph:*
44: gb_ph:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	968.6	81.3	1191	12 MMDHH	X76292 M.musculus
2	967.6	81.2	1190	5 AR021199	AR021199 Sequence
3	967.6	81.2	1190	5 AR063081	AR063081 Sequence
4	630.4	55.0	34325	51 AR011603	AR011603 Homo sapi
5	620.4	52.6	6206	9 AR010994	AR010994 Homo sapi
6	462	38.8	1958	4 GG058511	U58511 Gallus gall
7	429.4	36.1	2103	12 MMU85610	U85610 Mus musculus
8	419.8	35.2	1313	5 AR021201	AR021201 Sequence
9	419.8	35.2	1313	5 AR063083	AR063083 Sequence
10	418.2	35.1	1314	12 MMSHH	X76290 M.musculus
11	415.6	34.9	1197	4 XU026350	U26350 Xenopus lae
12	407	34.2	1715	12 RATVHH1X	L27340 Rat (Vh-1)
13	406.6	34.1	1191	4 XU026349	U26349 Xenopus lae
14	390.6	32.8	1425	5 AR063085	AR063085 Sequence
15	390.6	32.8	1576	9 HUMSHH	L38518 Homo sapien
16	385.2	32.3	1635	4 AR047466	AR047466 Sequence
17	381.4	32.0	1277	5 AR021198	AR021198 Sequence
18	381.4	32.0	1277	5 AR063080	AR063080 Sequence
19	381.4	32.0	1567	4 CHKZPAMED	L28099 Gallus gall
20	373	31.3	1718	4 AB029748	AB029748 Parolicht
21	370.2	31.1	1807	4 NEMSPA	D63339 Cynops pyr
22	357.4	30.0	1011	12 MKTHH	X76291 M.musculus
23	355.8	29.9	1056	5 AR021200	AR021200 Sequence
24	355.8	29.9	1056	5 AR063082	AR063082 Sequence
25	350.6	29.4	1549	4 AF003532	AF003532 Pleurodel
26	335.6	28.2	939	5 AR063086	AR063086 Sequence
27	335.6	28.2	1277	9 HUMTHH	L38517 Homo sapien
28	335.6	28.2	1277	13 G28584	G28584 human STS S
29	316.4	26.6	1251	4 DR030710	U30710 Danio rerio
30	316.2	26.5	1256	5 AR021202	AR021202 Sequence
31	316.2	26.5	1256	5 AR063084	AR063084 Sequence
32	316.2	26.5	1684	4 BRSONHP	L23569 B.rerio shh
33	316.2	26.5	2600	4 ZEPVHH1A	L27585 Brechydantio
34	314.6	26.4	1257	4 DR030711	U30711 Danio rerio
35	308.8	25.9	2447	4 XELXSH	L39213 Xenopus lae
36	308.8	25.9	386	12 AF148226	AF148226 Rattus no
37	307.2	25.8	1335	4 XU026314	U26314 Xenopus lae
38	305.2	25.6	1230	4 XU026404	U26404 Xenopus lae
39	304	25.5	1479	4 DRECHTDA	Y08426 D.rerio MRN
40	301.8	25.3	1491	9 AB018401	AB018401 Homo sapi
41	301.4	25.3	303	9 AB010581	AB010581 Homo sapi
42	301.4	25.3	2305	34 S66384	S66384 hh-segment
43	301.4	25.3	3091	34 DROHHA	L02793 Drosophila
44	299.8	25.2	2046	34 DROHHA	L05404 Drosophila
45	297.8	25.0	1546	4 XELVHH1A	L35248 Xenopus lae

ALIGNMENTS

	RESULT	1	
LOCUS	MMDH	1191 bp	DNA
DEFINITION	M.musculus (129/Sv) Dhh gene.		ROD
ACCESSION	X76292		18-JAN-1994
VERSION	X76292.1	GI:443941	
KEYWORDS	desert hedgehog protein; dhh gene.		
SOURCE	house mouse.		
ORGANISM	Mus musculus.		
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrate; Mammalia; Eutheria;		
AUTHORS	Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	McMahon,A.P.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (24-NOV-1993) A.P. McMahon, Harvard University, 16		
AUTHORS	Divinity Ave., Cambridge, MA 02138, USA		
TITLE	2 (bases 1 to 1191)		
JOURNAL	McMahon,J.A., Epstein,D.J., St-Jacques,B., Shen,L., Mohler,T.,		
FEATURES	Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity		
CELL LINE	Cell 75 (7), 1417-1430 (1993)		
LOCATION	Location/Ouallifiers		
source	1..1191		
/organism="Mus musculus"			
/strain="129/Sv"			
/db_xref="taxon:10090"			
/clone_lib="Stratagene; lambda fixit"			
sig-peptide	1..24		
/gene="Dhh"			
CDS	1..1191		
/gene="Dhh"			
/codon_start=1			
/product="desert hedgehog"			
/protein_id="CAA53924.1"			
/db_xref="GI:443942"			
/db_xref="MGD:MGI:94891"			
/db_xref="SPRMBL:O61488"			
/translation="MALPASILPLCLALLALSAOSGPGRPVRRARYRKOLVPLL			
KYKQPSMPERTLGASGAEGRVTRSGSRPFDLPVNPNDIIIFDENSGADRMTERR			
CKEYNVALAIVVMWGPVRRTREGMDSLEHGRLDITTSGRDNKYG			
LRLALAYAGCDWYTVESRNHIHSVDNLSTAVRAGCCPNATVLRSEGRGLAE			
LHRDWVLADPADRARVVFTPYLLFLDRDLORASFVAVEIERPKLLTPWLHVFPA			
RGPAPGEFAPVARRLRADGSVLAEGDMLQPARVARAREAAVGVPALTAHGTLL			
LVNVLASCAYVLESHOWMARAFAPLRLHLAIGLIPGANVOPTGMHYSLRIYLAEE			
ELMG"			
gene	1..1191		
/gene="Dhh"			
mat_peptide	22..1188		
/gene="Dhh"			
/product="desert hedgehog"			
BASE COUNT	195 a 371 c 399 g 226 t		
ORIGIN			
Query Match	81.3%	Score 968.6:	DB 12; Length 1191;
Best Local Similarity	88.3%:	Pred. No. 5.4e-153;	
Matches 1052; Conservative	0;	Mismatches 139;	Indels 0; Gaps 0;
Oy	1 ATGGCTCTCCTACCACACTACTGCGCTTGTCGTCGCACTTGTGAGCGTGCACGCC	60	
Db	1 ATGGCTCTGCGCGCCAGTCTGTGTCCTCTGCTTGCGCACTTTGGCACTATCTGCC	60	
Oy	61 CAGAGCTCGGGGCGGGGCGGGGCGGTTGGCCGCGCCCGCTATGCGCGCAAGCAGCTC	120	
Db	61 CAGAGCTCGGGGCGGGGCGGAGACCGGTTGGCCGCGGCGGCTTATGTGCGCAAGCACTT	120	
Oy	121 GTGGCGGCACTCTACAACGAATTGTTGGCGCGCGCGCGAGACGCAACCTTGGGCGGAGT	180	
Db	121 GTGGCTGTGCTATACAGAGATTGTGTGGCCAATATGACCGAGCGGAGCCTTGGGCGGAGT	180	

[illegible]

[illegible]

QY	781	CTTCACGCAAACTGTTGCTACAGCCGCTTGAGACACCGTGGTTTGGCGCTGAGAGGCGGAGCG	840
QY	781	CTTCACGCAAACTGTTGCTACAGCCGCTTGAGACACCGTGGTTTGGCGCTGAGAGGCGGAGCG <td>840</td>	840
Db	781	CTTCGCGCAAACTGTTGCTACAGCCGCTTGAGACACCGTGGTTTGGCGCTGAGAGGCGGAGCG	840
QY	841	CCCGCGCCAGAGGGAATTTGACACCGGTGTTCGGGCGCCCGCTACGCGCTGGGGACTCGGTG	900
Db	841	CTGCTCCAGAGGGAATTTGACACCGGTGTTCGGGCGCCCGCTACGCGCTGGGGACTCGGTG	900
QY	901	CTGGCGCCCGCGGGGAGATGCGCTTGGCCAGCGCGCGCTGGGCCGCTGTGGCGCGGAGAA	960
Db	901	CTGGCGCCCGCGGGGAGAGCGCGCTCCAGCGCGCGCGCTGAGCGCGCTGGCGCGGAGAA	960
QY	961	GGCGGGGGGGTTTGGCGCGCGCTACCGCGGACGGGAGAGCTGCTGTGAGATGATTCGTG	1020
Db	961	GGCGGGGGGGTTTGGCGCGCGCTACCGCGGACGGGAGAGCTGCTGTGAGATGATTCGTG	1020
QY	1021	GGCTTGTGCTACGCGGTTCGTGAGAGTACACAGTGGGCGCACCGCGCTTTTGGCCCTTG	1080
Db	1021	GGCTTGTGCTACGCGGTTCGTGAGAGTACACAGTGGGCGCACCGCGCTTTTGGCCCTTG	1080
QY	1081	AACACTGCTGACAGCGCTTAGGGGCGCTGCTCCCGCGCGCGCGCTACCGCGACTGCGATG	1140
Db	1081	CGGCTGCTGACAGCGCTTAGGGGCGCTGCTCCCGCGCGCGCGCTACCGCGACTGCGATG	1140
QY	1141	CATTGTACTTCGCGCTCCTTACCGCGCTTACGGGAGAGAGTACTGCGGTG	1190
Db	1141	CATTGTACTTCGCGCTCCTTACCGCGCTTACGGGAGAGAGTACTGCGGTG	1190
RESULT	3		
AR063081			
LOCUS	AR063081	1190 bp	DNA
DEFINITION	Sequence 2 from patent US 5844079.		PAT
ACCESSION	AR063081		
VERSION	AR063081.1	GI:5990772	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1190)		
AUTHORS	Ingham, P.W., McMahon, A.P. and Tabin, C.J.		
TITLE	Vertebrate embryonic pattern-inducing proteins, and uses related thereto		
JOURNAL	Patent: US 5844079-A 2 01-DEC-1998;		
FEATURES	Location/Qualifiers		
source	1..1190		
BASE COUNT	194 a 371 c 399 g 226 t		
ORIGIN			
Query Match	81.2%;	Score 967.6;	DB 5;
Best Local Similarity	88.3%;	Pred. No. 7.9e-153;	Length 1190;
Matches 1051;	Conservative	0;	Mismatches 139;
		Indels	0;
		Gaps	0;
QY	1	ATGGCTCTCTGACCAATCTACTGCGCTTGTGCTTGGACACTTCTGGCGCTGCACGC	60
Db	1	ATGGCTCTCTGACCGCCAGTCTGTGCCCTGTGCTCTGTGGCACTCTTGGCACTATCTGCC	60
QY	61	CAGAGCTGCGGGCGGGCGGGCGGGCGGGTGTGGCGGCGCGCGCTATGCGCGCAAGAGCTC	120
Db	61	CAGAGCTGCGGGCGGGCGGGCGGGCGGGTGTGGCGGCGCGCGCTATGCGCGCAAGACTT	120
QY	121	GTGGCGCTACTTACAGCAATTTGTGCCCGGCGTGGCAGAGCGGACCTTGGGCGCGAGT	180
Db	121	GTGGCTCTCTACTACAGCAAGTGTGTGCCCAATATGCCGACCGGACCTTGGGCGCGAGT	180
QY	181	GGGCGCAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	240
Db	181	GGGCGCAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	240
QY	241	TACACCCCGACATCATCTTCAAGGATGAGGAAACAGTGAAGCCGACCGCGCTGATGAC	300

Db	241	TACAA	CCCCCGCATATATCTTCAAGAGATGAGAGAACAGCGCGCGCAGACCGCCTGATGACA	300
QY	301	GAGCGTT	GCAAGAGAGAGGGTGAACGCTTTTGCCATTGCGCGTATGAACATGTGCGCCGGA	360
Db	301	GAGGGTT	GCAAAAGAGCGGGTGAACCGCTCTAGCCATGCGCGTATGAACATGTGCGCCGGA	360
QY	361	GTCGCGCT	CACAGAGTATGAGGGCTGGAGAGAGAGCGGCACACAGCTAGGATTCATCTC	420
Db	361	GTACGCC	TACTGTAATGAAAGCTGGAGACGAGAGAGCGGCACACGACACGACAGATTCATCTC	420
QY	421	CACATACGA	AGGCGCTGCTTTGGACATCTACTAGCTTGACCCGCGACCGCAACAATATGGG	480
Db	421	CACATACGA	AGGCGCTGCTTGGACATCTACTAGCTTGACCCGCGACCGCAACAATATGGT	480
QY	481	TTGTGGG	CGGCGCTCGAGATGGAAGCGCGGTTGATGAGTGGGTACTAGATGATCGCGCAAC	540
Db	481	TTGTGGG	CGGCGCTCGAGATGGAAGCGCGGTTGATGAGTGGGTACTAGATGATCGCGCAAC	540
QY	541	CACGCTCC	AGCTGTGATCAAAAGCTGATATATCACTAGCGCGGTCCGGCGCGGCGCTGCTTT	600
Db	541	CACATCCAC	GTATGATGATCAAAAGCTGATATATCACTAGCGCGGTCCGAGCGGAGGCTGCTTT	600
QY	601	CCGGGA	AAATGCACTGTGCGCCTGTGTGAGCGCGGACCGGAAGAGGCTCGCGGACTGCAC	660
Db	601	CCGGGA	AAATGCACTGTGCGCCTGTGTGAGCGCGGGAAGAGGCGCTGAGGGAACTACAT	660
QY	661	CGCGAG	AGATGGGTTTGGGGCGCGATAGCGCTCAAGCGCGGGTGGTGCCAGCGCGGTGCTG	720
Db	661	CGTGTG	ATCTGGTACTGAGTACGTGCGCGCTGATGACAGCGCGGCGGAGTGGTACCCAGCCATGCTG	720
QY	721	CTCTTC	CTGAGACCGGACTTGACAGCGCGCGGCTTCACTTTGTGGCTGTGAGACCGAGTGG	780
Db	721	CTCTTC	CTGAGACCGGAGATGACAGCGCGCGGCTTCACTTTGTGGCTGTGAGACCGAGTGG	780
QY	781	CCTCCAG	CAGCAACTGTGCTCAGCGCCCTGGCACCTGCTGTGTTGCGCGTGAAGGCGCGCG	840
Db	781	CCTCCG	CGCAAACTGTGCTCAGCACTCTGGCACTGTGTGTGCTGCTGCGCGGCGACCG	840
QY	841	CCCGCG	CACAGCGACTTGACACCGGTTTGCGCGCGCGGCTACGCGCTAGGAGACTCGGTCG	900
Db	841	CTCTCT	CCAGGTACTTGTGCACCGCGTGTGGCGCGCGGCTAGCTGCTGCGCACTCGGTCG	900
QY	901	CTGCGC	CGCGCGGAGTGCCTTGCGGCACGCGCGGTGGCCCGCTGTGCGCGCGGAGGAA	960
Db	901	CTGCGC	CGCGCGGAGGAGACGCGCTCACAACCGCGCGGTGAGCCCGCTGTGCGCGGAGGAA	960
QY	961	GCGCGG	CGCGGCTTCCGCGCGGCTCACCGCGGACGGAAGCACTGTGTGAACAATCTCTCG	1020
Db	961	GCGCGG	CGCGGCTTCCGCGCGGCTCACCGCGGACGGAAGCACTGTGTGAACAACCTCTCTC	1020
QY	1021	GCGCTT	GTCTACAGCGGCTTGAGAGTACACACAGATGGGCGCACCGGCTTTTCCGCCCTTGG	1080
Db	1021	GCGCTT	GTCTACAGCGGCTTGAGAGTACACACAGATGGGCGCACCGGCTTTTCCGCCCTTGG	1080
QY	1081	AGACTG	CTGACAGCGGCTTAGGGGCGCTGCTCCCGCGGGGCGGTCCAGCCGACTGGCATG	1140
Db	1081	CGGCTG	CTGACAGCGGCTTAGGGGCGCTGCTCCCGGGGGTGCAGTCCAGCCGACTGGCATG	1140
QY	1141	CATTGT	ACTCTGGGCTGCTCTACACGCTTAGGGGAGAGAGCTATCTGGGCGTG	1190
Db	1141	CATTGT	ACTCTGGCTCTCTTACACGCTTAGGGGAGAGAGTATCTGGGCGTG	1190

RESULT	4	
AC011603		
LOCUS	AC011603	241375 bp DNA
DEFINITION	Homo sapiens clone Rpl1-386G11, WORKING DRAFT SEQUENCE, 89	19-FEB-2000
ACCESSION	AC011603	
VERSION	AC011603.8	GI:7007704
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.	
SOURCE	human.	

ORGANISM

REFERENCE

AUTHORS

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Homo sapiens
 Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Mammalia:
 Eutheria: Primates: Catarrhini: Hominoidea: Homo.
 1 (bases 1 to 241275)
 Muzny,D.M., Adams,C., Bailey,M., Barberia,J., Blankenburg,K.,
 Bodada,B., Bouck,J., Bowle,S., Brooks,A., Buha,C., Bunac,C.,
 Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
 David,R., Delgado,O., Deshpande,D., Ding,Y., Domah-Rashid,N.,
 Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraruto,D.,
 Forcurn-Tansey,J., Frantz,F., Ganesh,R., Gorell,J.H., Gorell,L.L.,
 Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,
 Hollaway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
 Kelly,S., Kondajewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,
 Licharev,O., Liu,J., Liu,W., Logan,O., Lozano,R.J., Lu,Y.,
 Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M.,
 Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S.,
 Owsal,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L.,
 Quiles,M., Reiter,D., Rivers,M., Samuel,S., Say,J., Scherer,S.,
 Shah,E., Shen,H., Simon,M., Sparks,A., Stamps,A., Suggang,R.,
 Taber,P., Taylor,T., Vasquez,L., Vinson,R., Vo,Q., Wabhad,N.,
 Watlington,S., Weinstein,G., Weinstein,I.R., Williamson,A.,
 Worley,K., Wren,J., Wrenford,G., Yu,W., Zhou,X., Nelson,D., and
 Gibbs,R.
 Direct Submission
 Unpublished
 2 (bases 1 to 241275)
 Worley,K.
 Direct Submission
 Submitted (08-OCT-1999) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Feb 19, 2000 this sequence version replaced gi:6728920.

 Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

 Project Information
 Center project name: HMLM
 Center clone name: RP11-386G11

 Summary Statistics
 Sequencing vector: M13: L08821
 Chemistry: Dye-primer: Body: 2% of reads
 Chemistry: Dye-terminator: Big Dye: 98% of reads
 Assembly program: Phrap: version 0.980611
 Consensus quality: 166379 bases at least Q40
 Consensus quality: 189167 bases at least Q30
 Consensus quality: 200038 bases at least Q20
 Estimated insert size: 16846; agarose-1p estimation
 Estimated insert size: 20764; sum-of-contrigs estimation
 Quality coverage: 3.4x in Q20 bases; agarose-1p estimation
 Quality coverage: 3.1x in Q20 bases; sum-of-contrigs estimation

 NOTE: This is a 'working draft' sequence. It currently
 consists of 89 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

 1
 775: contig of 775 bp in length
 776 795: gap of unknown length
 796 1619: contig of 824 bp in length
 1620 1639: gap of unknown length
 1640 2534: contig of 885 bp in length
 2535 2554: gap of unknown length
 2555 3411: contig of 857 bp in length
 3412 3431: gap of unknown length
 3432 4198: contig of 767 bp in length
 4199 4218: gap of unknown length
 4219 5149: contig of 931 bp in length
 5150 5169: gap of unknown length

```
* 5170 5969: contig of 800 bp in length
* 5970 5989: gap of unknown length
* 5990 5987: contig of 889 bp in length
* 6879 6898: gap of unknown length
* 6899 7680: contig of 782 bp in length
* 7681 7700: gap of unknown length
* 8799 8798: contig of 1098 bp in length
* 8819 8818: gap of unknown length
* 9847 9846: contig of 1028 bp in length
* 9867 9866: gap of unknown length
* 10688 10687: contig of 821 bp in length
* 10708 10707: gap of unknown length
* 11476 11475: contig of 768 bp in length
* 11496 11495: gap of unknown length
* 12711 12710: contig of 1215 bp in length
* 12731 12730: gap of unknown length
* 13677 13676: contig of 946 bp in length
* 13697 13696: gap of unknown length
* 14553 14552: contig of 856 bp in length
* 14573 14572: gap of unknown length
* 15550 15549: contig of 977 bp in length
* 15570 15569: gap of unknown length
* 16388 16387: gap of unknown length
* 16408 16407: gap of unknown length
* 17183 17182: contig of 775 bp in length
* 17203 17202: gap of unknown length
* 18113 18112: contig of 910 bp in length
* 18133 18132: gap of unknown length
* 18952 18951: contig of 820 bp in length
* 18973 18972: gap of unknown length
* 19674 19673: contig of 701 bp in length
* 19694 19693: gap of unknown length
* 20611 20610: contig of 917 bp in length
* 20631 20630: gap of unknown length
* 21485 21484: contig of 855 bp in length
* 21506 21505: gap of unknown length
* 22337 22336: contig of 832 bp in length
* 22358 22357: gap of unknown length
* 23123 23122: contig of 766 bp in length
* 23143 23142: gap of unknown length
* 23960 23959: contig of 817 bp in length
* 23981 23980: gap of unknown length
* 25140 25139: contig of 1160 bp in length
* 25161 25160: gap of unknown length
* 26665 26664: contig of 1505 bp in length
* 26686 26685: gap of unknown length
* 28361 28360: contig of 1676 bp in length
* 28382 28381: gap of unknown length
* 29305 29304: contig of 924 bp in length
* 29326 29325: gap of unknown length
* 30969 30968: contig of 1643 bp in length
* 30989 30988: gap of unknown length
* 32141 32140: contig of 1153 bp in length
* 32162 32161: gap of unknown length
* 33671 33670: contig of 1510 bp in length
* 33692 33691: gap of unknown length
* 34584 34583: contig of 893 bp in length
* 34604 34603: gap of unknown length
* 35544 35543: contig of 940 bp in length
* 35565 35564: gap of unknown length
* 36680 36679: contig of 1096 bp in length
* 36681 36680: gap of unknown length
* 37453 37452: contig of 773 bp in length
* 37473 37472: gap of unknown length
* 38798 38797: contig of 1325 bp in length
* 38819 38818: gap of unknown length
* 39597 39596: contig of 779 bp in length
* 39617 39616: gap of unknown length
* 40293 40292: contig of 676 bp in length
* 40313 40312: gap of unknown length
* 41739 41738: contig of 1426 bp in length
* 41759 41758: gap of unknown length
* 42693 42692: contig of 934 bp in length
```

```
* 42694 42713: gap of unknown length
* 42714 43579: contig of 866 bp in length
* 43580 43599: gap of unknown length
* 43600 44425: contig of 826 bp in length
* 44426 44445: gap of unknown length
* 44446 45613: contig of 1168 bp in length
* 45614 45633: gap of unknown length
* 45634 47217: contig of 1584 bp in length
* 47218 47237: gap of unknown length
* 48026 48025: contig of 789 bp in length
* 48027 48046: gap of unknown length
* 48047 49628: contig of 1582 bp in length
* 49629 49648: gap of unknown length
* 49649 50758: contig of 1110 bp in length
* 50759 50778: gap of unknown length
* 50779 51801: contig of 1023 bp in length
* 51802 51821: gap of unknown length
* 51822 53254: contig of 1433 bp in length
* 53255 53274: gap of unknown length
* 53275 54149: contig of 875 bp in length
* 54150 54169: gap of unknown length
* 54170 55527: contig of 1358 bp in length
* 55528 55547: gap of unknown length
* 55548 56345: contig of 798 bp in length
* 56346 56365: gap of unknown length
* 56366 57710: contig of 1345 bp in length
* 57711 57730: gap of unknown length
* 57731 59144: contig of 1414 bp in length
* 59145 59164: gap of unknown length
* 59165 60516: contig of 1352 bp in length
* 60517 60536: gap of unknown length
* 60537 61275: contig of 739 bp in length
* 61276 61295: gap of unknown length
* 61296 62780: contig of 1485 bp in length
* 62781 62800: gap of unknown length
* 62801 64415: contig of 1615 bp in length
* 64416 64435: gap of unknown length
* 64436 65722: contig of 1287 bp in length
* 65723 65742: gap of unknown length
* 65743 67483: contig of 1741 bp in length
* 67484 67503: gap of unknown length
* 67504 69118: contig of 1615 bp in length
* 69119 69138: gap of unknown length
* 69139 71052: contig of 1914 bp in length
* 71053 71072: gap of unknown length
* 71073 72873: contig of 1801 bp in length
* 72874 72893: gap of unknown length

Query Match 52.9% Score 630.4; DB 57; Length 241275;
Best Local Similarity 99.1%; Pred. No. 5.7e-97;
Matches 634; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 552 GTCGCTCAAGCTGATACACTGCGCGTCCGGCGCGCGCTTCCGGGAATGC 611
    |||
DB 99285 GTTGTCCTCAAGATACACTGCGCGTCCGGCGCGCGCTTCCGGGAATGC 99324

QY 612 AACGTGCGCTGTGAGCGCGGAGCGGAAGGCTGCGGGAATGCACCCGGAAGCTG 671
    |||
DB 99325 AACGTGCGCTGTGAGCGCGGAGCGGAAGGCTGCGGGAATGCACCCGGAAGCTG 99384

QY 672 GATTGTGCGCGCATGCGTGCAGCGCGGCTGCCACGCGGCTGCTCTTCCTGGA 731
    |||
DB 99385 GATTGTGCGCGCATGCGTGCAGCGCGGCTGCCACGCGGCTGCTCTTCCTGGA 99444

QY 732 CCGGACTTGCAGCGCGCGGCTTCAATTGCTGTGAGACGAGTGCTCCACGCAA 791
    |||
DB 99445 CCGGACTTGCAGCGCGCGGCTTCAATTGCTGTGAGACGAGTGCTCCACGCAA 99504

QY 792 ACTGTGCTACGCGCTGTGAGCGCGGCTTGTGCGGCTGGAAGCGCGCGCGCGCCAGG 851
    |||
DB 99505 ACTGTGCTACGCGCTGTGAGCGCGGCTTGTGCGGCTGGAAGCGCGCGCGCGCCAGG 99564

QY 852 GCATTTCACCGGCTGTGCGCGCGCGCTACGCGGCTGCGGACTCGGCTGTGCGCGCGG 911
```

DB	99565	99566	99567	99568	99569	99570	99571	99572	99573	99574	99575	99576	99577	99578	99579	99580	99581	99582	99583	99584	99585	99586	99587	99588	99589	99590	99591	99592	99593	99594	99595	99596	99597	99598	99599	99600	99601	99602	99603	99604	99605	99606	99607	99608	99609	99610	99611	99612	99613	99614	99615	99616	99617	99618	99619	99620	99621	99622	99623	99624	99625	99626	99627	99628	99629	99630	99631	99632	99633	99634	99635	99636	99637	99638	99639	99640	99641	99642	99643	99644	99645	99646	99647	99648	99649	99650	99651	99652	99653	99654	99655	99656	99657	99658	99659	99660	99661	99662	99663	99664	99665	99666	99667	99668	99669	99670	99671	99672	99673	99674	99675	99676	99677	99678	99679	99680	99681	99682	99683	99684	99685	99686	99687	99688	99689	99690	99691	99692	99693	99694	99695	99696	99697	99698	99699	99700	99701	99702	99703	99704	99705	99706	99707	99708	99709	99710	99711	99712	99713	99714	99715	99716	99717	99718	99719	99720	99721	99722	99723	99724	99725	99726	99727	99728	99729	99730	99731	99732	99733	99734	99735	99736	99737	99738	99739	99740	99741	99742	99743	99744	99745	99746	99747	99748	99749	99750	99751	99752	99753	99754	99755	99756	99757	99758	99759	99760	99761	99762	99763	99764	99765	99766	99767	99768	99769	99770	99771	99772	99773	99774	99775	99776	99777	99778	99779	99780	99781	99782	99783	99784	99785	99786	99787	99788	99789	99790	99791	99792	99793	99794	99795	99796	99797	99798	99799	99800	99801	99802	99803	99804	99805	99806	99807	99808	99809	99810	99811	99812	99813	99814	99815	99816	99817	99818	99819	99820	99821	99822	99823	99824	99825	99826	99827	99828	99829	99830	99831	99832	99833	99834	99835	99836	99837	99838	99839	99840	99841	99842	99843	99844	99845	99846	99847	99848	99849	99850	99851	99852	99853	99854	99855	99856	99857	99858	99859	99860	99861	99862	99863	99864	99865	99866	99867	99868	99869	99870	99871	99872	99873	99874	99875	99876	99877	99878	99879	99880	99881	99882	99883	99884	99885	99886	99887	99888	99889	99890	99891	99892	99893	99894	99895	99896	99897	99898	99899	99900	99901	99902	99903	99904	99905	99906	99907	99908	99909	99910	99911	99912	99913	99914	99915	99916	99917	99918	99919	99920	99921	99922	99923	99924	99925	99926	99927	99928	99929	99930	99931	99932	99933	99934	99935
----	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

Query Match Similarity 52.6%; Score 626; DB 9; Length 626;
Best Local Similarity 100.0%; Pred. No. 1e-95;
Matches 626; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```

OY 566 ATMACTCACTGGCGGTCCGGGGCGCGCGCTGCTTTCCGGGAAATGCAACTGTGCGCTGT 625
DB 1 ATAACTCACTAGCGGTCCGGGGCGGGGGGTGCTTTCCGGGAAATGCAACTGTGCGCTGT 60
OY 626 GGAGCGCGGAGCGGGAAGGGCTGCGGGAACTCACCCGGGACATCTGGTTTGGCGGGC 685
DB 61 GGAGCGCGGAGCGGGAAGGGCTGCGGGAACTCACCCGGGACATCTGGTTTGGCGGGC 120
OY 686 ATGCGTCAAGCGCGGGTGTGTCACGCGCGCTGCTGCTTTCCGTGACCGGAGCTTGCAG 745
DB 121 ATGCGTCAAGCGCGGGTGTGTCACGCGCGCTGCTGCTTTCCGTGACCGGAGCTTGCAG 180
OY 746 GCCGGGCTTCAATTTGTGCTGTGAGACCGAATGGCTTCCAGCAAACTGTTGCTACGC 805
DB 181 GCCGGGCTTCAATTTGTGCTGTGAGACCGAATGGCTTCCAGCAAACTGTTGCTACGC 240
OY 806 CTTGGCACCTGTGTGTTGGCGCTGACGAGCGCGCGCGCGCGCGAGCTTGTGACAGG 865
DB 241 CCTGGCACCTGTGTGTTGGCGCTGACGAGCGCGCGCGCGCGAGCTTGTGACAGG 300
OY 866 TGTTCGCGCGCGCGCTACCGCGCTGGGGACTGCTGTGCGCGCGCGCGGGAGTGCCTTC 925
DB 301 TGTTCGCGCGCGCGCTACCGCGCTGGGGACTGCTGTGCGCGCGCGCGGGAGTGCCTTC 360
OY 926 GCGCAGCGCGCGCTGGCGCGCTGTGGCGGGGAGGAAGCGTGGGCGTGTCCGCGCGCTCA 985
DB 361 GCGCAGCGCGCGCTGGCGCGCTGTGGCGGGGAGGAAGCGTGGGCGTGTCCGCGCGCTCA 420
OY 986 CCGCGCACGGAGAGCGTGTGCTGTGAACGAATGCTCTGCTTGTGCTACGCGGTTCTGAGA 1045
DB 421 CCGCGCACGGAGAGCGTGTGCTGTGAACGAATGCTCTGCTTGTGCTACGCGGTTCTGAGA 480
OY 1046 GTCAACAGTGGGCGACACGCGCTTTTGGCCCCCTTGAGACTGCTGCACGCGCTAAGGGCGC 1105
DB 481 GTCAACAGTGGGCGACACGCGCTTTTGGCCCCCTTGAGACTGCTGCACGCGCTAAGGGCGC 540
OY 1106 TGCCTCCCGGCGGGGCGCGTGCACGCGACTGGATGGATGTGATCTGCGCTCCTTACC 1165
DB 541 TGCCTCCCGGCGGGGCGCGTGCACGCGACTGGATGGATGTGATCTGCGCTCCTTACC 600
OY 1166 GCTTAGCGGAGAGCTACTGAGGCTGA 1191
DB 601 GCTTAGCGGAGAGCTACTGAGGCTGA 626

RESULT 6
LOCUS GGU58511 1958 bp mRNA VRT 16-AUG-1996
DEFINITION Gallus gallus indian hedgehog protein mRNA, complete cds.
ACCESSION U58511
VERSION U58511.1 GI:1493841
KEYWORDS
SOURCE
ORGANISM
  chicken.
  Gallus gallus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria;
  Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
  1 (bases 1 to 1958)
REFERENCE
  Vortkamp,A., Lee,K., Lanske,B., Segre,G.V., Kronenberg,H.M. and
  Tabin,C.J.
  Regulation of rate of cartilage differentiation by Indian hedgehog
  and PTH-related protein
  Science 273 (5275), 613-622 (1996)
  2 (bases 1 to 1958)
  Vortkamp,A., Lee,K., Lanske,B., Segre,G.V., Kronenberg,H.M. and
  Tabin,C.J.
  Direct Submission

```


QY	521	TCCTACAGAGCGCCGACCAACAGCGTCACAGTGTGGGTGTAAGAGCTATACCTACACGGGG	580
Db	521	TCCTACTAGATCCAAAGACTCATCATCTGTTCTGTAAAGCAAGAACTCCGTGGGG	580
QY	581	TCCGGGCGGGGCGTGCCTTTCCTGGGAAATGCAACTGTGCGCTGTGTGAACGGCGAGCGGA	640
Db	581	CCAAATCCGGGGGCGTGTTCCTCCGGGATCCGGACCGTGCACCTTGAAGAGCGGCGACCA	640
QY	641	AAGGCTCCGGGAACATGCACCGCGGAGACTGGGTTTGGCGGGCGATCGTCAGGCCGAG	700
Db	641	AGCTGTGTAAGGACTTACGTCCCGGAGACCGGTGCTGTGCGGCTGACGACCAAGGCGCGC	700
QY	701	TGGGCCCCAGCCGGTGTGCTCTTCGAGAACCGGAGACTTGAAGCGCGGGCTTATTG	760
Db	701	TGCTGTACAGGACTTCTCTACCTCTTCGTGAGCCGCGAGAGAGCGCCCAAGAAAGTCTTCT	760
QY	761	TGGCTGTGAGAGCCAGTGGGCTCCACGCAAACTGTTGTCTACGCTTGAGCACTGTGT	820
Db	761	ACGTGATGAGACGCTGGAGCGCGCGAGAGCGCTGTCTGTCTACCGCGGACCTGTCT	820
QY	821	TTGCCGCTC-----GAGGGCGGGCCCGCGCCAGCGACTTTGCACCGGT	868
Db	821	TGTGGGCCCGCACACACTCGGGGGCCACGGCCCGGCGCAAGCGCTCTTTGCGACGC	880
QY	869	TGGCGCGCGGCTACGGGCTGGGGAGCTGGGTGTGGCCCGCGCGGGGA---TGGGCTTC	925
Db	861	GGTTCGCGCCCGGGGAGCGCGCTGTACTGTGTGTGTAACGCGCGCGGGAGCGCGCGGTGC	940
QY	926	GGCCAGCGCGGTGGCCCGCTGTGGC---GCGGAGAGAAAGCCGTGGCGGTATCGCGCGC	982
Db	941	TGCCGCGCGGTTGCACAGCGTGAAGCTGAGAGAGAGAGAGAGGCGGGCGGTACGGCGCGC	1000
QY	983	TCACCGCGCAGCGGACGCTGCTGTGTGAACGATGTCTTGCGCTCTTGCTACGCGGTTTGG	1042
Db	1001	TCACGCGCCACGCGGCACATTCATCAACAACCGGGTGTCTGCGCTACGCTGTCACTG	1060
QY	1043	AGAGTCACAGAGGGGCGGACCGCGGCTTTTGCCCTTGAGATGTGTCACAGCGCTAGGGG	1102
Db	1061	AGGACACAGCTGGGCGACACCGGCGCTTCGCGCTTTCCGCTTGCGCGCACGCGCTGTGG	1120
QY	1103	CGCTGCT 1109	
Db	1121	CCGCGCT 1127	
RESULT	9		
LOCUS	AR063083	1313 bp	DNA
DEFINITION	Sequence 4 from patent US 5844079.	PAT	29-SEP-1999
ACCESSION	AR063083		
VERSION	AR063083.1	GI:5990774	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1313)		
AUTHORS	Ingham,P.W., McMahon,A.P. and Tabin,C.J.		
TITLE	Vertebrate embryonic pattern-inducing proteins, and uses related thereto		
JOURNAL	Patent: US 5844079-A 4 01-DEC-1998;		
FEATURES	Location/Qualifiers		
source	1..1313		
BASE COUNT	254 a 411 c 428 g 220 t		
ORIGIN	/organism="unknown"		

Query Match	35.28;	Score 419.8;	DB 5;	Length 1313;
Best Local Similarity	64.58;	Pred. No. 2.3e-61;		
Matches 701; Conservative	0;	Mismatches 362;	Indels 24;	Gaps 4

```

OY      41 CACTTCTGCGCTGCGACCCAGAGCTGCGCGGCCGGGGCCGGGTTGCGCGGGCC 100
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

[illegible]

RESULT 11
LOCUS XLU26350 1197 bp mRNA VRT 05-JAN-1996
DEFINITION Xenopus laevis hedgehog protein 4 (hh4) mRNA, complete cds.
ACCESSION U26350
VERSION U26350.1 GI:1147829
KEYWORDS
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae; Xenopodinae;
Xenopus.
REFERENCE
AUTHORS Ekker, S.C., McGrew, L.L., Lai, C.J., Lee, J.J., von Kessler, D.P.,
Moon, R.T. and Beachy, P.A.
TITLE Distinct expression and shared activities of members of the
hedgehog gene family of Xenopus laevis
JOURNAL Development 121 (8), 2337-2347 (1995)
MEDLINE 95401852
REFERENCE 2 (bases 1 to 1197)
AUTHORS Ekker, S.C., McGrew, L.L., Lai, C.J., Lee, J.J., von Kessler, D.P.,
Moon, R.T. and Beachy, P.A.
TITLE Direct Submission
SUBMITTED (04-MAY-1995) Stephen C. Ekker, Molecular Biology and
Genetics, Johns Hopkins University, 725 N. Wolfe St/714 PCTB,
Baltimore, MD 21205, USA
FEATURES
source location/Qualifiers
1..1197
/organism="Xenopus laevis"
/db_xref="taxon:8355"
1..1197
/gene="hh4"
1..1197
/note="related to Xenopus laevis cephalic hedgehog,
encoded by Genbank Accession Number U26349; secreted
signaling molecule; Method: conceptual translation
supplied by author"
/codon_start=1
/product="hedgehog protein 4"
/protein_id="AA85164.1"
/db_xref="GI:1147830"
/translation="MPAVRILILAACCCWLLLPVRCGPGRGVGGRRRMRRLVPL
LYKQFVNPYPEKTLGASGSEKIRGSEKILVNYNNDILFKDEENYGAARLMT
RCDKRVNALAISVMNMPGLKLVTEGMDGHANDSLHREBALDITTSDDRKRY
GMARLAVAGAGFDMVYSEKAIHVSUNVDNSGVRSGGCFPTAVMMETGKKRLS
ELKIDPTVTTTIDETGLIHISVULFLHROPYKATVYLLEABHPKRLVTPNHLPT
KSSSTGFQPTFAIVQIDGLDQIYNGTVOVSSKVRVSVDSQTVAPMTEHGLL
VGVLTSCYATVESHVTLAHASLAPRLFLFOGIASMLPDLHSDVHWYCHILYLAKYV
LWMDMP"
BASE COUNT 295 a 286 c 314 g 302 t
ORIGIN
Query Match 34.3%; Score 415.6; DB 4; Length 1197;
Best Local Similarity 62.2%; Pred. No. 1.2e-60;
Matches 653; Conservative 0; Mismatches 409; Indels 12; Gaps 2;
Db 4 GCCTCTGACCAATCTACGCTTGTGCTGCTTGGCAGCTTCGCGGCGCCAG 63
7 GCACTCCGAGTTCTGATACAGCTGCTGCTGCTGCTGCTTCCCTCCAGTTCCG 66
QY 64 AGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 123
Db 67 TGTGTGTGCTCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 126
QY 124 CCGCTACTACAGCAATTTGTGCGCGGCGTGCAGAGCGGAGCCCTGGGCGCAGTGGG 183
Db 127 CCACTGCTTACAGAGAGTTGTGCTCAATGTGCCAGAGAGACTAGAGGCGCAGTGGC 186
QY 184 CCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 243

Db 187 AACTCAGAGGGCAAGATCCGACGCGGCTTCGAAAGGTTTCATCAAGTTGCCCAACTAC 246
QY 244 AACCCGACATCATCTTCAAGAGATGAGAGAACAGTGAACCGGCGGCTGATACCGAG 303
Db 247 AACCCAGACATCATCTTCAAGAGATGAGAGAACAGAGAGAGAGAGAGAGAGAGAG 306
QY 304 CGTTGCAGAGAGAGGAGTGAACGCTTTGGCCATTGCCGTATGAACATGTGGCCGAGTG 363
Db 307 CGCTGCAAGACAGAGATCATGCTTGGCTATCTGTGTATGAACATGTGGCAGGCTTG 366
QY 364 CGCTGCAAGACAGAGATCATGCTTGGCTATCTGTGTATGAACATGTGGCAGGCTTG 423
Db 367 AACCTGCGGAGTACAGAGGCGGCTGGATGAGAGCGGCGACACGCTCAGCAGCTTTG 426
QY 424 TAGCAAGCGGCTCTTGTGACATCATCTGACATCTGACCGGAGCGGCAACAGTATGG 483
Db 427 TAGCAAGCGGAGCGGCTTGTGACATCATCTGACCGGAGCGGCAACAGTATGG 486
QY 484 CTGCGCGGCTGCGCATGAGAGCGGCGGCTTGTGACATCTGACCGGAGCGGCAAC 543
Db 487 TTGGCCGCTGCGCATGAGAGCGGCGGCTTGTGACATCTGACCGGAGCGGCAAC 546
QY 544 GTCCAGCTGCGGCTGCAAGCTGATTAATCTGACGCGGCTCCGCGGCGGCGGCTTCC 603
Db 547 ATCCAGCTGCGGCTGCAAGCTGATTAATCTGACGCGGCTCCGCGGCGGCGGCTTCC 606
QY 604 GGAAGTCACTGTGCGGCTGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 663
Db 607 GGAAGTCACTGTGCGGCTGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 666
QY 664 GGAAGTCACTGTGCGGCTGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 723
Db 667 GGTGACACTGTCTTCACTACATGAGATGAAGAGCGGCTGTGATATCATGAGTTCCT 726
QY 724 TTCTGAGCGGAGCTTGCAGCGGCGGCTTCAATTTGTGCTGCTGAGACCGAGTGCCT 783
Db 727 TTCTGAGCGGAGCTTGCAGCGGCGGCTTCAATTTGTGCTGCTGAGAGCGGAGTGC 786
QY 784 CCAAGCAACTGTGCTGCAAGCGGCGGCTTCAATTTGTGCTGCTGAGAGCGGAGTGC 843
Db 787 CCAAGCAACTGTGCTGCAAGCGGCGGCTTCAATTTGTGCTGCTGAGAGCGGAGTGC 846
QY 844 GCGCGAGGCGGCTTGCAGCGGCGGCTTGCAGCGGCGGCTTGCAGCGGCGGCTTGC 902
Db 838 TCATCAAGAGATTTCAACCACTTTGCTTATAGAGTTGTAATAGAGAGCTGATTC 897
QY 902 -TGGCGCGGCGGAGTGGCTTGCAGCGGCGGCTTGCAGCGGCGGCTTGCAGCGG 960
Db 898 ATCTATGTAATGGCACTCAGGTGCAATCTTCAAGGTTGTGAGGATTCCTGATGAA 957
QY 961 GCGGTGCGGCTTGCAGCGGCGGCTTGCAGCGGCGGCTTGCAGCGGCGGCTTGCAG 1020
Db 958 CCAAGTGAATGTAATGCTTCCATGACTGAGACATGGAACCTTATTTGATGGAGTCT 1017
QY 1021 GCCTCTGCTACGCGGCTTCTGAGAGTCAACAGTGGGCGGCGGCTTGTGCGGCTT 1080
Db 1018 ACATGCTGCTATGCTACCGGAGTCAACAGTGGGCGGCGGCTTGTGCGGCTT 1077
QY 1081 AGACTGCTGACGCGCTGAGGCGGCTTGTGCGGCTTGTGCGGCTTGTGCGGCTT 1114
Db 1078 AGGCTCTTCAAGGAGTCAAGTCTTCAAGTCTTCAAG 1111

RESULT 12
LOCUS RATVH1X 1715 bp mRNA ROD 25-AUG-1994
DEFINITION Rat (vhh-1) mRNA, complete cds.
ACCESSION L27340
VERSION L27340.1 GI:452122
KEYWORDS
ORGANISM Rattus norvegicus strain Sprague-Dawley CDNA to mRNA.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

QY	258	CTCAAGGATGATGGAGAAACACTGGAGCGACCCGCTGATATGACCGAGCGTTGCAAGGAGACG	317
Db	255	ATTTAAAGATGAGAAAAACACCGGAGGGGACAGGCTTATCTACGTACAGGGGTGTAAGACAA	314
QY	318	GGTGAACGCTTTGGCGCATTTGCCGTGATGATGAACATGTGGCCCGGAGTGGCGCTACAGATGAC	377
Db	315	GTTGAACGCTTTGGCGCATCTCGGTGATGAACCAATGTGGCCAGAGGTGAATCGCGGGGTGAC	374
QY	378	TGAGGGCTGGGACGACAGACGGCCACACGCTCAGAGATTCTCCACTACGAAAGCCGTGC	437
Db	375	CGAGGGCTGGGACGAGAAATGGCCACACACTAGAGAGACTCTCTCACTACGAGGGCCGCGC	434
QY	438	TTTGACATCTACTACGTGTGACCCCGACCCGACCAACATATGGGTGTGGCGCGCCCTGC	497
Db	435	AGTGAACATCTACACAGCTGTGACCCCGACCCGACGAGATACGGCATGTGCGCCCGCTGGC	494
QY	498	AGTGAAGCCGGCTTGCAGTGGGTCTACTACAGAGTCCCGAACCCAGCTCAGCTGTGGT	557
Db	495	GGTGAAGCCGGCTTGCAGTGGGTCTACTACAGATCCAGACCATATACATCTCACTGTCTGGT	554
QY	558	CAAACTCTATTAATCTACATGGGGGTCCGGGGGGGGGGTGGCTTTCCGGGAAATTCATCTGT	617
Db	555	GAAGACAGAGAACTCGGTGGGGGGGCAAAATGGGAGGTCTGTTCCCGAGGTCTGGCCACGCT	614
QY	618	GGCGCTGTGGAGCGCGGACGCGGAAAGGGTGTGGGAACTCCACCGCGGAAATCTGGTTTT	677
Db	615	GCACCTGGAGGACGAGGGCGGACCAACAGCTGTGTAAAGACTTAGGCCCGGGGAGCCGCTGCT	674
QY	678	GGCGGCGCATCGTCGTAGCGCGGGGTGTGTGCCACGCCGCGTCTCTTCTGTGACCGGGA	737
Db	675	GGCGGCGGACGACACAGGCGCGGGCTGTCTACAGGAGATTCCTCACTTCTGTGACCGGGA	734
QY	738	CTTGGAGGGCGCGGCTTATTTGTGGCTGTGAGAGACCGAATGGCCCTCCACGCAAACTGTT	797
Db	735	CGACGGCGCCCAAGAAAGTCTTCTACGTGATCGAAGCGGGGAGCCCGCGAGCCGCTGCT	794
QY	798	GCTCACGCCCTGGGACCTGTGTGTTGCCGC-----TCGAGGGCCGCGCGCCCGC	846
Db	795	GCTCACGCCCGCGGACCTGTGTGTTGGCGCGCACAAAGACTCGGCACCGGGGAGGCC	854
QY	847	CCAGCGCACTTTGCACCGGTGTTGGCGCGCGGCTACGCGCTGGGGACTGGTGTGCGG	906
Db	855	CGAGCGCTCTCGGGCTCTGGGGCGCGCTTCCGGGGGCGCACTAGGGGCTTGGGCGCTGTT	914
QY	907	CCCGCGGGGGAATGGCGCTTCGGCCACAGCGGGGTGGCCCGTG-----	947
Db	915	CGCCACGCCGCTGCGCCCGGGGCCACGCGCTGTGTACGTGTGGCCGACGTGACGGGACCG	974
QY	947	-----TGCGCGGGAGGAAGCCGTGGCGCTGTT	974
Db	975	CCGGCTCTCTGCCCGCGCGCTGTGACACAGGTGACCTTAAGGAGAGAGCGCGGGGCGCTTA	1033
QY	975	CGCGCGCGCTCACCGCGGACGGGAGAGCTGTGTGTGAACAGATGTCTTGAGCTCTTGTCTACGC	1033
Db	1035	CGCGCGCGCTCACCGCGGACCGGACCACTTCTATCAACCGGGGTCTGTGCTGTGTGTACGC	1094
QY	1035	GGTTCTGAGAGATCCACAGTGGGGGACCGGGGCTTTTGGCCCTTGAGTGTGTGTACGC	1094
Db	1095	GGTCAATCAAGAGACACAGCTGGGGCGACCGGGGCTTTGGCGCCCTTCCGCTGTGGCGCACGC	1155
QY	1095	GCTAGGGGC 1103	
Db	1155	GCTCTGCG 1163	
RESULT	15		
LOCUS	HUMSHH	1576 bp	mrna
DEFINITION	Homo sapiens sonic hedgehog protein (SHH) mRNA, complete cds.		
ACCESSION	L38518		
VERSION	L38518.1	GI:66156	
KEYWORDS	homologue; sonic hedgehog protein.		

SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES
CDNA to mRNA.	Homo sapiens (clone HHH5)	(bases 1 to 1576)	Marigo, V., Roberts, D.J., Lee, S.M.K., Tsukunov, O., Levy, T., Copeland, N.G., Seidman, D.J., Gilbert, D.J., Martin, G.G., McMahon, A.P., and Tabin, C.	Cloning, expression and chromosomal location of SHH and IHH, two human homologues of the Drosophila segment polarity gene Hedgehog unpublished (1995)		
Location/Qualifiers						
1..1576	/organism="Homo sapiens"					
/db_xref="taxon:9606"						
/clone="HHH5"						
/dev_stage="fetus"						
/tissue_type="lung"						
/tissue_lib="Clontech"						
/note="vector: lambda gt10"						
1..1576	/partial					
/gene="SHH"						
1..1576	/gene="SHH"					
152..1540	/gene="SHH"					
/gene="SHH"						
/note="homologue of the Drosophila segment polarity gene"						
/codon_start=1						
/product="sonic hedgehog protein"						
/protein_id="AA62179.1"						
/db_xref="GI:663157"						
/translation="MLLRRLCLLVVSLVCSGLACGPGRGKRRHPKLLPLAYKQFLPAEKTIGASGRYEGKISNSRPFELPNYPDIIIFKRENTGADRLMTORCKDKNALAISYMNQMPGVKRLRYTGMDMDGHSESHYSGRADYITSDRDRKTYGM LARIAPAGFDWVYTESKAHHCVRKAENYAAASGGCFESATVHLEGGTKLVKDL SPGRVRLVADDOGRLLTFLDRDQKAKVEYVETREPRRLILLTAHLFVAP HNSSEAGAVASGSGSPSGALGPRLAFASRVAPQGVVVARDEDRLPLPAVHS VTSLEAAGVAPLTAOGTILINRVLASCAVIEHSMARHAFAPFLAHLAALAP ARDPRGDSGGGDRGGGGRVALPAPADAPAGATAGIHWTSQLLYIGTWLDSSE ALHPDGAAYKSS"						
BASE COUNT	287 a	508 c	567 g	214 t		
ORIGIN						
Query Match	32.8%	Score 390.6	DB 9	Length 1576		
Best Local Similarity	61.9%	Pred. No. 1.7e-56				
Matches 711: Conservative	0	Mismatches 369	Indels 69	Gaps 3		
18	1CTACTGCCCTGTGCTGCTCTGCTGACCTTCTGAGCGCTGCGACCCAGAGCTCGGCGCGG	77				
DB	172 TCTGCTGCTAGTCCCTGCTCTCTGCTGCTGTAAGTCTGGGAGACGTGGCGTGGACCGG	231				
78	CCGGGCGCGGTTGGCGCGCGCGCTATGCGCGCAGCAGCTGTCGCGCTACTCTACAA	137				
DB	232 CAGGGGCTTCGGGAGAGAGAGGAC-----CCAAAAGCTGACCCCTTAGCTACAA	285				
138	GCAATTTTGTGCGCGCGCTGCCAGAGCGGACCTTGGGCGCCAGTGGGCCAGGAGGGAG	197				
DB	286 GCAGTTTATCCCAATGTGGCGGAAACCCCTAGCGCCACGCGGAAAGTATGAAGGGA	345				
198	GGTGCAAGGGGCTCCGAGCGCTTCGCGGAGCTGTCGCGCAACTTACACCCCGACATAT	257				
DB	346 GATCTCCGAAACTCGAGCGATTTAAGGAACTACCCCAATTTACACCCCGACATAT	405				
258	CTTCAAGATGAGAGACAGTGAAGCCGACCGCTGATGACCGAGCGCTTCAAGAGAG	317				
DB	406 ATTTAAGATGAGAGAAACCGGACCGGACAGCGTATGACTCAGAGTGTAAAGCAA	465				
318	GGTGAACGCTTTGGCCATTTGCCGTGATGAACATGTGGCCCGAGTGGCGCTACGAGTAC	377				

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2000, 06:53:49 ; Search time 52.13 Seconds
(without alignments)
4044.423 Million cell updates/sec

Title: US-08-900-220-7

Perfect score: 1622
Sequence: 1 CATCACCCACGAGAGACC.....CATTTGGAGGCCCATTCCTCC 1622

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 460926

Minimum DB seq length: 0
Maximum DB seq length: 100000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

Issued Patents, NA: *
1: /cgn2_6/ptodata/1/1na/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/1na/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/1na/5C.COMB.seq: *
4: /cgn2_6/ptodata/1/1na/5D.COMB.seq: *
5: /cgn2_6/ptodata/1/1na/6.COMB.seq: *
6: /cgn2_6/ptodata/1/1na/PTCUS.COMB.seq: *
7: /cgn2_6/ptodata/1/1na/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	936	57.7	939	3	US-08-356-060A-7 Sequence 7, Appli
2	843.2	52.0	1056	2	US-08-176-427B-5 Sequence 5, Appli
3	843.2	52.0	1056	3	US-08-356-060A-3 Sequence 3, Appli
4	459	28.3	1277	2	US-08-176-427B-1 Sequence 1, Appli
5	459	28.3	1277	3	US-08-356-060A-1 Sequence 1, Appli
6	453.8	28.0	1425	3	US-08-356-060A-6 Sequence 6, Appli
7	452.2	27.9	1576	2	US-08-748-591-5 Sequence 5, Appli
8	452.2	27.9	1576	2	US-08-748-591-10 Sequence 10, Appli
9	438.8	27.1	1313	2	US-08-176-427B-7 Sequence 7, Appli
10	438.8	27.1	1313	3	US-08-356-060A-4 Sequence 4, Appli
11	432	26.6	1715	6	PCT-US95-02315-1 Sequence 1, Appli
12	419.8	25.9	1190	2	US-08-176-427B-3 Sequence 3, Appli
13	419.8	25.9	1190	3	US-08-356-060A-2 Sequence 2, Appli
14	418.2	25.8	1256	2	US-08-176-427B-9 Sequence 9, Appli
15	418.2	25.8	1256	2	US-08-356-060A-5 Sequence 5, Appli
16	144	8.9	144	6	PCT-US95-15463-2 Sequence 2, Appli
17	144	8.9	144	6	PCT-US95-15463-1 Sequence 1, Appli
18	112	6.9	144	6	PCT-US95-15463-1 Sequence 1, Appli
19	112	6.9	144	6	PCT-US95-15463-1 Sequence 1, Appli
20	47.6	2.9	2887	5	US-09-183-253-1 Sequence 1, Appli
21	47	2.9	44377	3	US-08-804-227C-7 Sequence 7, Appli
22	47	2.9	44377	4	US-08-804-198-1 Sequence 1, Appli
23	44.8	2.8	1433	4	US-08-666-392A-1 Sequence 1, Appli
24	43.8	2.7	765	5	US-08-718-904-79 Sequence 79, Appli
25	43	2.7	4524	2	US-08-845-998-7 Sequence 7, Appli
26	42.8	2.6	2004	2	US-08-471-033-18 Sequence 18, Appli
27	42.8	2.6	2004	3	US-08-471-044-18 Sequence 18, Appli

28	42.8	2.6	2004	3	US-08-463-483A-18 Sequence 18, Appli
29	42.8	2.6	2004	3	US-08-471-046A-18 Sequence 18, Appli
30	42.8	2.6	2004	3	US-08-470-566B-18 Sequence 18, Appli
31	42.8	2.6	2004	4	US-08-469-334-18 Sequence 35, Appli
32	42.8	2.6	2576	2	US-08-471-033-35 Sequence 35, Appli
33	42.8	2.6	2576	3	US-08-471-044-35 Sequence 35, Appli
34	42.8	2.6	2576	3	US-08-463-483A-35 Sequence 35, Appli
35	42.8	2.6	2576	3	US-08-470-566B-35 Sequence 35, Appli
36	42.8	2.6	2576	4	US-08-469-334-35 Sequence 35, Appli
37	42.8	2.6	2655	2	US-08-471-033-17 Sequence 17, Appli
38	42.8	2.6	2655	2	US-08-471-033-26 Sequence 26, Appli
39	42.8	2.6	2655	3	US-08-471-044-17 Sequence 17, Appli
40	42.8	2.6	2655	3	US-08-471-044-26 Sequence 26, Appli
41	42.8	2.6	2655	3	US-08-463-483A-17 Sequence 17, Appli
42	42.8	2.6	2655	3	US-08-463-483A-26 Sequence 26, Appli
43	42.8	2.6	2655	3	US-08-471-046A-17 Sequence 17, Appli
44	42.8	2.6	2655	3	US-08-471-046A-26 Sequence 26, Appli
45	42.8	2.6	2655	3	US-08-471-046A-26 Sequence 26, Appli

ALIGNMENTS

RESULT 1
US-08-356-060A-7
Sequence 7, Application US/08356060A
Patent No. 5844079
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tablin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
Proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHYE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356, 060A
FILING DATE: 14-DEC-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176, 427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMI-006CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 939 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..939
US-08-356-060A-7

Query Match 57.7%; Score 936; DB 3; Length 939;
 Best Local Similarity 100.0%; Pred. No. 1.7e-198;
 Matches 936; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 351 CGCCTATGACCCAGCCGCTGACAGGACCGCTGAACTCGCTGCTATCTCGGTGATGTAAC 410
DB 4 CGCCTATGACCCAGCCGCTGACAGGACCGCTGAACTCGCTGCTATCTCGGTGATGTAAC 63
QY 411 CAGTGGCCCGGTGAGAGCTGCGGGTACCGAGAGGCTGGAGACGAGAGCCACTCA 470
DB 64 CAGTGGCCCGGTGAGAGCTGCGGGTACCGAGAGGCTGGAGAGGAGAGCCACTCA 123
QY 471 GAGAGTCCCTGATTAATGAGAGCCCGCGGTGACATCACACATATAGACCGGACCG 530
DB 124 GAGAGTCCCTGATTAATGAGAGCCCGCGGTGACATCACACATATAGACCGGACCG 183
QY 531 AATAGATGAGAGTGGCGGGCGCTTGGAGTGGAGGCGGCTTGTACTGGGTATTAC 590
DB 184 AATAGATGAGAGTGGCGGGCGCTTGGAGTGGAGGCGGCTTGTACTGGGTATTAC 243
QY 591 GAGTCAAGGCCCACTGATGCTATGCTCCGTCAGTCCGAGCACTCGCCGCAAGCAAGAC 650
DB 244 GAGTCAAGGCCCACTGATGCTATGCTCCGTCAGTCCGAGCACTCGCCGCAAGCAAGAC 303
QY 651 GCGGCGTCTTCCCTGCGGAGCCCAAGTACGCTTGAAGAGTGGGGCGCTGTGGCTTG 710
DB 304 GCGGCGTCTTCCCTGCGGAGCCCAAGTACGCTTGAAGAGTGGGGCGCTGTGGCTTG 363
QY 711 TCAGCCGTGAGCGCGGAGACCGTGTGCTGGCCATGGGGAGATGGAGACCCCACTTC 770
DB 364 TCAGCCGTGAGCGCGGAGACCGTGTGCTGGCCATGGGGAGATGGAGACCCCACTTC 423
QY 771 AGCATGTGCTATTTCTCTGAGACCGGAGCCCAAGGCTGAGAGGCTTCAAGGTATC 830
DB 424 AGCATGTGCTATTTCTCTGAGACCGGAGCCCAAGGCTGAGAGGCTTCAAGGTATC 483
QY 831 GAGACTAGAGACCCCAAGCGCGCTGCACTACACCGCTACCTGCTTTAGCGCT 890
DB 484 GAGACTAGAGACCCCAAGCGCGCTGCACTACACCGCTACCTGCTTTAGCGCT 543
QY 891 GACATACACGAGGAGCGGCGGCTTCCGGGCACTTTGGCAGACGATGACGCT 950
DB 544 GACATACACGAGGAGCGGCGGCTTCCGGGCACTTTGGCAGACGATGACGCT 603
QY 951 GCGCAGTACGTGTGTGCGGTGAGGCTGACAGGCTGACGCTGCGCGGTGACGCTGTC 1010
DB 604 GCGCAGTACGTGTGTGCGGTGAGGCTGACAGGCTGACGCTGCGCGGTGACGCTGTC 663
QY 1011 TCTACACAGTGGCCCTGCGGGGCTACGCGCGCTCACAAAGCATGGGACATGGTGTG 1070
DB 664 TCTACACAGTGGCCCTGCGGGGCTACGCGCGCTCACAAAGCATGGGACATGGTGTG 723
QY 1071 GAGATGTGGGATGCTGCTTGGGGCGGTGATGACCAACCCGAGGATGGTGGC 1130
DB 724 GAGATGTGGGATGCTGCTTGGGGCGGTGATGACCAACCCGAGTGGTGGC 783
QY 1131 TTTGCGCCCTGAGACTCTTTCACAGCTTGGCATGGGAGAGTGGAGAGGCT 1190
DB 784 TTTGCGCCCTGAGACTCTTTCACAGCTTGGCATGGGAGAGTGGAGAGGCT 843
QY 1191 GTGATTTGTAACCCCAAGCTGCTTACCGCGCTGGGGCGTCTCTGTAGAAAGGAGC 1250
DB 844 GTGATTTGTAACCCCAAGCTGCTTACCGCGCTGGGGCGTCTCTGTAGAAAGGAGC 903
QY 1251 TTTCAACCACTGGGATGCTCGGGGCAAGAGCTGA 1286
DB 904 TTTCAACCACTGGGATGCTCGGGGCAAGAGCTGA 939

```

RESULT 2
 US-08-176-427B-5
 ; Sequence 5, Application US/08176427B
 ; Patent No. 5789543

```

? GENERAL INFORMATION:
? APPLICANT: Ingham, Phillip W.
? APPLICANT: McMahon, Andrew P.
? APPLICANT: Tablin, Clifford J.
? TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
? TITLE OF INVENTION: Proteins and Uses Related Thereto
? NUMBER OF SEQUENCES: 33
? CORRESPONDENCE ADDRESS:
? ADDRESS: LAHIVE & COCKFIELD
? STREET: 60 State Street
? CITY: Boston
? STATE: MA
? COUNTRY: USA
? ZIP: 02109
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: ASCII(text)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/176,427B
? FILING DATE: 30-DEC-1993
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Vincent, Matthew P.
? REGISTRATION NUMBER: 36,709
? REFERENCE/DOCKET NUMBER: HMI-006
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (617) 227-7400
? TELEFAX: (617) 227-5941
? INFORMATION FOR SEQ ID NO: 5:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1056 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: both
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 1..1056
? US-08-176-427B-5

Query Match 52.0%; Score 843.2; DB 2; Length 1056;
Best Local Similarity 88.2%; Pred. No. 5.5e-178;
Matches 928; Conservative 0; Mismatches 123; Indels 1; Gaps 1;

```

Db 361 GCCGCTGCCAAGACAGCTGCTGCTTCTTCCGCGAGCCCAAGTGGCCCTAGAGAAAGG 420
QY 696 GCGCGTGTGCGCTTGTACAGCCCGTGAAGCCGAGACCGTGTGTGGCCATGGGGAGAT 755
Db 421 GAGCGTGTGCGCTGTACAGCTGTAAAGCCAGAGACCGGGTGTGCGCATGGGGAGAT 480
QY 756 GGGAGCCCACTTACAGCATGTGTCTATTTCTTGTGAGCCGAGCCCAAGAGCTGAGA 815
Db 481 GGGAGCCCACTTACAGCATGTGTCTATTTCTTGTGAGCCGAGCCCAAGAGCTGAGA 540
QY 816 GCGCTTCAGGTATGAGAGCTGAGACCCCGCGCGCTGTGCACTACACCGGCTGAC 875
Db 541 GCTTTCAGGTATGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 600
QY 876 CTGCTCTTACAGGTATGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 935
Db 601 CTGCTCTTACAGGTATGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 660
QY 936 AGCCAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 995
Db 661 AGCCAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 720
QY 996 GCGGTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1055
Db 721 GCGGTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 780
QY 1056 GGGAGACAGTGTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1115
Db 781 GGGAGACAGTGTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 840
QY 1116 CTGCTCTTACAGGTATGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 1175
Db 841 CTGCTCTTACAGGTATGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 900
QY 1176 ACCCGGGGAGAGTGTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 1235
Db 901 ACCCGGGGAGAGTGTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 960
QY 1236 CTGAGAGAGAGAGTGTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 1295
Db 961 CTGAGAGAGAGAGTGTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 1020
QY 1295 CACCGCTGCGCTTCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 1326
Db 1021 AACCACTGCGCTTCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 1052

RESULT 3
US-08-356-060A-3
Sequence 3, Application 08/08356060A
Patent No. 5844079
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tabin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356, 060A
FILING DATE: 14-DEC-1994

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMI-006CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1056 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1056
US-08-356-060A-3

Query Match 52.0%; Score 843.2; DB 3; Length 1056;
Best Local Similarity 88.2%; Pred. No. 5,5e-178;
Matches 928; Conservative 0; Mismatches 123; Indels 1; Gaps 1;

QY 276 GAGCGTTTAAAGAGTCAACCCCAATTCATCCAGACATATCTTCAAGAGCAGAG 335
Db 1 GAGCGTTTAAAGAGTCAACCCCAATTCATCCAGACATATCTTCAAGAGCAGAG 60
QY 336 AACACAGGCGCCAGCGCTCATGACCCAGCGCTGCAAGAGACCGCTGAACTGCTGCT 395
Db 61 AACACAGGCGCCAGCGCTCATGACCCAGCGCTGCAAGAGACCGCTGAACTGCTGCT 120
QY 396 ATCTGCTATGAAACAGTGGCGCGGTGTGAACCTGCGGTGACCGAGCGCTGGAGAG 455
Db 121 ATCTGCTATGAAACAGTGGCGCGGTGTGAACCTGCGGTGACCGAGCGCGGTGAA 180
QY 456 GAGCGCACACCTCAAGAGAGTCCCTGCAATATGAGAGCGCGCGGTGAGACATACACA 515
Db 181 GATGCTATGAAACAGTGGCGCGGTGTGAACCTGCGGTGACCGAGCGCGGTGAA 240
QY 516 TCAGACCGGAGCGCAATAGTATGAGAGTGGCGCGGTGAGAGCGCGGTGAGAGCG 575
Db 241 TCAGACCGGAGCGCAATAGTATGAGAGTGGCGCGGTGAGAGCGCGGTGAGAGCG 300
QY 576 GACTGGGTATTTAAGAGTCAAGAGCCCAAGTGCATTTGCTCCGTCAGTCCGAGACTCG 635
Db 301 GACTGGGTATTTAAGAGTCAAGAGCCCAAGTGCATTTGCTCCGTCAGTCCGAGACTCG 360
QY 636 GCGCGACCAAGAGCGCGCGGTCTTCCCTGCGGAGCCAGAGTACGCTGAGAGTGG 695
Db 361 GCGCGACCAAGAGCGCGCGGTCTTCCCTGCGGAGCCAGAGTACGCTGAGAGTGG 420
QY 696 GCGCGTGTGCGCTTGTGAGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 755
Db 421 GAGCGTGTGCGCTTGTGAGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 756 GGGAGCCCACTTACAGCATGTGTCTATTTCTTGTGAGCCGAGCCCAAGAGCTGAGA 815
Db 481 GGGAGCCCACTTACAGCATGTGTCTATTTCTTGTGAGCCGAGCCCAAGAGCTGAGA 540
QY 816 GCGTTCAGGTATGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 875
Db 541 GCGTTCAGGTATGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 600
QY 876 CTGCTCTTACAGGTATGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 935
Db 601 CTGCTCTTACAGGTATGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 660
QY 936 AGCCAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 995

Db	661	AGCCATGTGCACCAAGGCCAATAATGTCGTCTCAGAGGGGTACAGAGGCGCTCCAGCGTCT	720
Oy	986	CGCGTGCGACGTGTCTCTAACACACGTGTGCCCCCTAGGCCCGCTCACAAAGCAT	10555
Db	721	CGGGTGGAGTGTCTCTCCACCACAGTGGCCCTTTGGGTCTCTATGTCTCTTCCAAAGGAT	780
Oy	1056	GGAGACTGGTGTGTGGAGATGTGTGGATCTCGCTTGGCGGCGCTGGCTACCAACAC	1115
Db	781	GGGACACTGTGTGTGGAGATGTGTGTGGCTCTCTGTGTGGACGTGTGGCTATACCAACAT	840
Oy	1116	CTGCGCTCAGTTGGCTTCTGGCCCCCTGAGACTCTTTCACAAGCTTGGCATGGGCGAGCTGG	1175
Db	841	CTGGCTCAGTTGGGCTCTCTGGCCCCCTGGCACTTTTCCAGTTTGGCATGGGCGAGCTGG	900
Oy	1176	ACCCGGGGGGGGGTGTGCAATTGGTAACCCGACCTGTCTACCGCGCTGGGGGCTACTCGT	1235
Db	901	ACCCCAAGTAGGGTGTCTTACTCTCTTACCTCTCAATGCTCTACCGCGCTGGGGGCTCTTGG	960
Oy	1236	CTAGAAGAGGGGACACTTCCACCACACTGTGGCATGTCCGGGGCAGAGAGCTGAAGAGCATC-	1295
Db	961	CTAGAAGAGAGACACTTCCATCCACTGTGGGCGATGTCTGGGGCAGAGAGCTGAAGAGCATCT	1020
Oy	1295	CACGCGTGCCTCTCTGGAAGTGTCTGTACTGTGGG	1336
Db	1021	AACCACTGTCCCTCTGTGAACTGTGTGTGCTGGG	1052

```

1      RESULT      4
2      US-08-176-427B-1
3      ; Sequence 1, Application US/08176427B
4      ; Patent No. 5789543
5      ;
6      ; GENERAL INFORMATION:
7      ; APPLICANT: Ingham, Phillip W.
8      ; APPLICANT: McMahon, Andrew P.
9      ; APPLICANT: Tabin, Clifford J.
10     ; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
11     ; TITLE OF INVENTION: Proteins and Uses Related Thereto
12     ; NUMBER OF SEQUENCES: 33
13     ; CORRESPONDENCE ADDRESS:
14     ; ADDRESSEE: LAHIVE & COCKFIELD
15     ; STREET: 60 State Street
16     ; CITY: Boston
17     ; STATE: MA
18     ; COUNTRY: USA
19     ; ZIP: 02109
20     ;
21     ; COMPUTER READABLE FORM:
22     ; MEDIUM TYPE: Floppy disk
23     ; COMPUTER: IBM PC compatible
24     ; OPERATING SYSTEM: PC-DOS/MS-DOS
25     ; SOFTWARE: ASCII(text)
26     ;
27     ; CURRENT APPLICATION DATA:
28     ; APPLICATION NUMBER: US/08/176,427B
29     ; FILING DATE: 30-DEC-1993
30     ; CLASSIFICATION: 435
31     ;
32     ; ATTORNEY/AGENT INFORMATION:
33     ; NAME: Vincent, Matthew P.
34     ; REGISTRATION NUMBER: 36,709
35     ; REFERENCE/DOCKET NUMBER: HMT-006
36     ;
37     ; TELECOMMUNICATION INFORMATION:
38     ; TELEPHONE: (617) 227-7400
39     ; TELEFAX: (617) 227-5941
40     ;
41     ; INFORMATION FOR SEQ. ID NO.: 1:
42     ;
43     ; SEQUENCE CHARACTERISTICS:
44     ; LENGTH: 1277 base pairs
45     ; TYPE: nucleic acid
46     ; STRANDEDNESS: both
47     ; TOPOLOGY: linear
48     ; MOLECULE TYPE: cDNA
49     ; FEATURE:
50     ;
51     ; NAME/KEY: CDS
52     ; LOCATION: 1..1277
53     ;
54     ; US-08-176-427B-1

```

Query Match	28.38;	Score 459;	DB 2;	Length 1277;
Best Local Similarity	65.28;	Pred. NO. 3.9e-93;		
Matches 763;	Conservative	0;	Mismatches 360;	Indels 48;
			Gaps	4;

QY	163	GGCAGCCGCCAACCACAACTGTCGCGCTGCTACAAACAGTTCAGCCCAATGTGCGC	222
Db	107	GGAGGACCCCAAAAAAGCTGACCCCGTTAGCCTATATACAGATTATATCCAAATGTGGAC	166
QY	223	AGAAAGCCCTGGCGCCACGAGACGCTATGAAAGCAAGTGCCTCGACGTCCAGCGCT	282
Db	167	AGAAAGACCCCTAGGGGCGCAATGGAAAGTATGAAGGGGAAGATCAACAAGAAATCCCAAGAT	226
QY	283	TCAAAGAGCTCACCCCAATTACAAATCCAGACATATCTTCACAGACAGAGAAACAG	342
Db	227	TTAAAGAACTAACCCCAATTACAAACCTGCACATATATTTTAAGATATAAAGAAACAG	286
QY	343	GCGCCGACGCGCTCATATGACCCACAGCGCTGCAGAGACCGCGCTGAATCGCTGCTATCGG	402
Db	287	GAGCTGCACAACTGATGACTCAGCGCTGCAGAGACAACTGMAATGCCGTGCAGATCTCG	346
QY	403	TGATGACACAGTGGCGCGGATGTAACCTCGCGGTGATCCGAGGGCGGGACAGACAGCC	462
Db	347	TGATGAACCAAGTGGCGCGGGGTGTAAGCTGCGGGTATCCGAGGGCGCTGGAGACAGAGATGGC	406
QY	463	ACCACCTCAGAGAGTCCCTGCATTTATGAGCGCGCGGTGATCCACATCAGATCAGACC	522
Db	407	ATCATCTCCGAGGATGCTGTGCATCTACAGAGGGTGGCGCGGTGACATCCACAGTCCGATC	466
QY	523	GCGACCCCATTAAGTATGAGACTGCTGGCGCGGTTGSCAGTGGAGGCGCGCTTACATGGG	582
Db	467	GGGACCCGACCAAGTACGGAATCTGTGGCCGCTGCGCGTGAAGCGCGGCTTGCACATGGG	526
QY	583	TGATTTACAGATCAAAAGCCCAAGTGCATTTGCTCCGTCAAGTCCGAGCATCTGGCGCAG	642
Db	527	TCTACTACGAGTCCAAAGCGCACATCCATGCTCCGTCGTAAGACAGAAAACTCACTGGCAG	586
QY	643	CCAAAGAGGGGCGGCTCTTCCGCGCGGAGGCCAGTACGCTGGAGATGGGGCGGGTG	702
Db	587	CGAATTAGAGGCTGCTTCCCTGCTCAGCCACAGTGCATCTGGAGATGAGAGCACCA	646
QY	703	TGCGCTTGTACCGCTGAGAGCGCGGAGACCGTGTGCTGGCATGGGGAGAGATGGAGCC	762
Db	647	AGCTGGTGAAGAGCACTGAGCGCCGCGGAGCCGCGTGTGCTGATGACGAGACGGCGGC	706
QY	763	CCACCTTACAGGATGTGCTCATTTTCTGTGGACCGGAGCGCCACAGGCTGAGACCTTCC	822
Db	707	TGCTCTACATGACTCTCTCATCTCTCTGTGACCGGATGACAGCTCCCGAAGCTCTTCT	766
QY	823	AGGTCACTGAGACTTCAGAGACCCCCACAGCGCGCTGGCACTACATCCGCTACCTGCTCT	882
Db	767	ACGTCACTGAGAGCGCGGACGCCCGGCGCCGCGTCTACTGACGGCGGCCACACTGCTCT	826
QY	883	TTTACGCTGACATCATCACGAGACCGG-----CAGCCGCTCTCCGGGCGCACAT	930
Db	827	TTTGGGCCCCCAGCAACCACTGTGGAGGCCACAGGGTCCACAGTGGCCAGGCGCTCT	886
QY	931	TTTGCACCCACAGTGCAGCTGGCCA-----GTACGTGCTGTGCTGTGGGTGCGCAGGCC	984
Db	887	TCGCGACCAACGTAAGCGCTGGCCAAACGTTGTTATGTGTGGCGCAGGGCGGACAGGC	946
QY	985	TGCAGCTGCGCGCGGTGGCAGCTGTCTTACAC---ACGTGGCCTCGGGGCTTACGCCCC	1041
Db	947	TGCTGCGCGGGGTGTGTCACAGGCTGTCATTCGCGGAGAGAGCGTCGCGACCTTACGCCC	1006
QY	1042	CGCTCAACAAGCAATGGGACACTGTTGGTGGAGATGTGGTGCATCTGCTCCGCGCG	1101
Db	1007	CACCTACCGGCCACGGGCAACATCTCATTAACCGGGGTGTGGCCCTCGTCAACGCCGCA	1066
QY	1102	TGCGTACCAACCACTGAGCTCAAGTGGCCTTGTGGCCCTGAGACTCTTTTACACAGTTGG	1151
Db	1067	TCGAGGAGACAGATTGGGCCATTTGGGCCCTTGGCACATTTCCGCTTGGCTACAGGGCTGC	1126

QY 1162 CATGGGCGAGCTGAGACCCCGGGGAG-----GGTGTGC 1194
Db 1127 TGGCGCCCTCTGGCCCAATGAGGCGATCCTACTGCGCCACACACCACTGGCATCC 1186
QY 1195 ATTGTATACCCAGCTGTCTACCGCCCTGGGGCGCTCTCTGCTAGAAAGGCGAGCTTCC 1254
Db 1187 ATGTGATCTACGCGCTCTCTACCGCATCGGAGCTGGGTCTGTGATGATGACGCGCTGC 1246
QY 1255 ACCACTGGGATGTCGGGGGAGGAGCTG 1285
Db 1247 ATCCGCTGGGATGTGTGACCGGCGACCTG 1277

RESULT 5

US-08-356-060A-1
Sequence 1, Application US/08356060A
Patent No. 5844079

GENERAL INFORMATION:

APPLICANT: Ingham, William W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tablin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA

ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,060A
FILING DATE: 14-DEC-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMT-006CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1277 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:

NAME/KEY: CDS
LOCATION: 1..1275
US-08-356-060A-1

Query Match 28.3%; Score 459; DB 3; Length 1277;

Best Local Similarity 65.2%; Prem. No. 3.9e-93;
Matches 763; Conservative 0; Mismatches 360; Indels 48; Gaps 4;

QY 163 GCGGACCCGACGCAAACTCGTCCGCTGCTTACAGAGAGTTACGCCCATGTGCCCG 222
Db 107 GGAGGACACCCCAAAAGCTAGACCCCGCTTAGCTATAGAGAGTTATCCCAATGTGGCAG 166
QY 223 AGAAGACCCCTGGGCGCCAGCGGACGCTATGAAGGAAGATCGCTGCAAGCTTCGAGCGCT 282

Db 167 AGAAGACCCCTAGGGGCGCCAGTGAAGATATGAAGGAGATACAAAGAACTCCAGAGAT 226
QY 283 TCAAGGAGCTACCCCGCAATATACATCCAGCATCATCTTCAAGAGAGAGAAACAG 342
Db 227 TTTAAGACTAACCCCAATATACACCTGTACATTTATTTAAAGATGAAGAACAGCGG 286
QY 343 GCGCGGACCGCTATGACCCAGCGCTGCAAGGACCCGCTTAACTCGCTGCTATCTCG 402
Db 287 GAGCTGACAGACTGATGACTAGCGCTGCAAGGACAAAGCTAATAGCCTGCGATCTCG 346
QY 403 TGATGAACTATGCGCGGCTGTAACTGCGGCTGACCGAGGCTGGAGAGAGAGCGC 462
Db 347 TGATGAACCACTGCGCGGCTGTAACTGCGGCTGACCGAGGCTGGAGAGAGATGCGC 406
QY 463 ACCACTGAGAGAGTCCCTGATATGAGGGCGCGGCTGGACATACCAATCAGACC 522
Db 407 ATCACTCCGAGGAATCGTGCATACGAGGCTCGCGCTGGACATCACAGTGGATC 466
QY 523 GCGACCGCAATATGATGAGCTGTGGCGCGCTTGGCAGTGGAGCGCGCTTGTACTGG 582
Db 467 GGGACCGGACGAAGTACGATGCTGTGGCGCGCTTGGCAGTGGAGCGCGCTTGTACTGG 526
QY 583 TGATTTACAGTCAAAAGGCGCCAGCTGATGCTCGCTGATGAGTCCGAGACTGCGCGCAG 642
Db 527 TCTACTAGAGTCCAAAGCGCCACATCTCCTCGTCAAGCAAGAACTCAGTGGCAG 586
QY 643 CCAAGAGCGGGGCTGCTTCCCTGCGCGAGTCCAGTACCGCTGAGAGTGGGCGCGCTG 702
Db 587 CGAATACGAGAGCTGCTTCCCTGCTGACGACAGTCCAGTGGAGATGAGAGACCA 646
QY 703 TGGCTGTGACCGCTGAGGCGCGGAGACCGTGTGCTGCTGAGGAGGAGATGAGAGC 762
Db 647 AGCTGTGAAGAGACTGAGCGCTGGGAGACCGGCTGTGCTGCTGAGAGCGGAGCGCGC 706
QY 763 CCACCTTACGAGTGTCTATTTTCTGAGACCGGAGCCCAAGCTGTGAGAGCTTCC 822
Db 707 TGCCTTACAGTACTTCTCCTCCTCCTGACCGAGTGAAGTCCCGAAAGCTTCT 766
QY 823 AGGTATGAGAGTCAAGAGACCGCGCAGCGCTGGGACCTACACCGCGCTACCGTCT 882
Db 767 ACGTATGAGAGCGGAGCGGAGCGCGGCGCGCTGTACTACAGCGCGCCACCTGCTCT 826
QY 883 TTACGCTGACAAATCACAGGAGCGCG-----CAAGCGCTTCCGGGCGACAT 930
Db 827 TTGTGGCCCCCAGACAAACAGTGGAGGCGACAGGCTCACAGTGGCGAGGGCTCT 886
QY 931 TTGCGAGCCAGTGCAGCTGCGCA-----GTACGCTGTGTGGCTGGGTGCCAGGCG 984
Db 887 TCGCCAGCAAGTGAAGCTGTGCAAGCTGTATGTGTGTGGGAGGAGGCGGCGAGCAGC 946
QY 985 TGCAGCTGCGCGGCTGCGACCTGTCTACAG---AGTGGCGCTCGGGCGCTACAGCGC 1041
Db 947 TGTGCGCGCGCTGTCTCACAAGCGCTCATTTGCGGAGAGAGCGTCCGAGCCTACGCGC 1006
QY 1042 CGCTTCAAAGCAGTGAAGTGTGTGAGAGTGTGTGTGAGTCTCTGCTTCCGCGCGC 1101
Db 1007 CACTACCGCGCCAGGAGCAACATCTCATCAACCGGCTTGTGCTCTGCTACAGCGCGTCA 1066
QY 1102 TGGCTGACCAACCACTGCTGCTAGTTGGCTTGTGGCCCTGAGACTCTTTACAGCTTGG 1161
Db 1067 TCGAGAGACACAGTGTGGCGCATTTGGCGCTTGCACCATCTCGCTTGGCTCAGGGGCTGC 1126
QY 1162 CATGGGCGAGTGAACCCCGGGGAG-----GGTGTGC 1194
Db 1127 TGGCGCCCTCTGGCCCAATGAGGCGCATCCTACTGCGCCACACCACTGGCATCC 1186
QY 1195 ATTGTATACCCAGCTGTCTACCGCCCTGGGGCGCTCTCTGCTAGAAAGGCGAGCTTCC 1254
Db 1187 ATGTGATCTACGCGCTCTCTACCGCATCGGAGCTGGGTCTGTGATGATGACGCGCTGC 1246
QY 1255 ACCACTGGGATGTCCGGGGAGGAGCTG 1285

Oy	390	CTGGCTATCTCGGGGATGAACCACTGGCCCGGGGTGAACCTGGGGGTGACCCAGGGGCTGG	449
Db	476	TTGGCCATCTCGGTGATTAACCACTGGCCAGAGTGAATCGGGGTGACCCAGGGGCTTGG	535
Oy	450	GACGAGACGGCCACCACTCAGAGGAATCCCTGCATTAATAGGGGCCGCCGGTGGACATC	509
Db	536	GACGAAATGAGCCACCACTCAGAGGAATCTCTGCACTACGAGGGGCCGCAATGGAGCATC	595
Oy	510	ACCAATATAGACCCCGACCCGCAATAAATATGAGACTCTGGCCGCTTGGCAATGGAGGCC	569
Db	596	ACCAAGTGTGACCCCGACCCGCAATAGGGAATGCGATGCTGGCCCGCTGGCGGTGGAGGCC	655
Oy	570	GACCTTGCATGGGTGTATTACGATCAAAAGGCCACAGTGTGATTGCTCCGTCAAGTCCAG	629
Db	656	GCGTTCCACTGGGTGTACTACGATCCAAAGCCAAATATCACTGCTCGGTGAAGACAGAG	715
Oy	630	CACTCGGCCGACGCCAAGACGGGGCGGTGCTTCCCTCCGCGAGCCCAAGGTACGCTCGAG	689
Db	716	AACTCGGTGGGGCCAAATGGGGAGGTGGTTCCCGGGGCTCGGCCACAGGTGCACTGGAG	775
Oy	680	AGTGGGGCGCTGTGGCTTGTCAGCCGTGAGGCCGGGAGACCGTGTGCTGGCCATGGGG	749
Db	776	CAGGGCCGACCAAGACTGTGTGAAGAACTGAGCCCCGGGGACCGCGTGTGGCGGCGGAC	835
Oy	750	GAGGATGGGAAACCCCACTTCAGAGATGTCTATTTTCTGTGACCCGCAAGCCCAACAG	809
Db	836	GACCAAGGGCCGGCTCTCTACAGGAATTTCTCACTTCTCTGAGCGCCACGAGAGGGCGC	895
Oy	810	CTGAGAGCCTTCCAGTCAATGAGACTCAGAGACCCCAACGCCGCGCTGCATCTCACACC	869
Db	896	AAGAAGGCTTCTACAGTGAATCGAAGCCGGGAGACCCGCGAGAGGCTCTGCTGTCAACGC	955
Oy	870	GCTACACTGCTCTTTACGGCTGACAAATCACAGGAGCCGAGACCCGCTTCGGGGCCACA	929
Db	956	GCGCACTGCTCTTTGTGGC---GCCGACAAAGAACTCGACCAACCGGAGGCCGAGGCG	1012
Oy	930	TTTCCGACGCAACGTGAGCCTGTGGCAATACGTCTCTGTGGTGGGGTGC	978
Db	1013	TCCTGGGCTCGGGGCGGCTTCCGGGGGGGCACTGAGGGCTCTGGGGCGC	1061

RESULT 9
 US-08-176-427B-7
 Sequence 7, Application US/08176427B
 Patent No. 5789543
 GENERAL INFORMATION:
 APPLICANT: Ingham, Phillip W.
 APPLICANT: McMahon, Andrew P.
 APPLICANT: Tablin, Clifford J.
 TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
 TITLE OF INVENTION: Proteins and Uses Related Thereto
 NUMBER OF SEQUENCES: 33
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LAHIVE & COCKFIELD
 STREET: 60 State Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII(text)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/176,427B
 FILING DATE: 30-DEC-1993
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: HMI-006

```

: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 227-7400
: TELEFAX: (617) 227-5941
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1313 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1314
:
US-08-176-427B-7

```

[illegible]

Db	280	GCGGACAGGCTCAGACACAGAGATCGAAGAAGACAACTGAACCTCGGTGGCATCTCTCTGA	339
QY	405	ATGACACAGTGGCCCCGGGTGTGAAGCTGGGGTACCGAGGGCTGGGACGAGGACGGCCAC	464
Db	340	ATGAACCACTGGCCACGGGGTTAAGACTCGGTGACAGAGAGGGCTGGGATGAGACGGTCAC	399
QY	465	CACTCAGAGAGTCCCTGTCATTATGAGGGCGCGCGGTGACATCAACCATCAACAGCC	524
Db	400	CATTTTGAAGATATCTCCACTACGAGAGGAAGCTGTGATATTAACCACTCTACCA	459
QY	525	GACCGCATATAGTATGGACAGCTGGCCGGCTTGGACAGTGAAGCGCGCTTACTGGGTG	584
Db	460	GACACAGCAATATCGGGACACTGTCTCCGCTTACGTGTGAAGGCTGATTTACTGGGTC	519
QY	585	TATTACGAGTCMAAGGCCCAAGTCATTTGCTCCGTCAAGTCGAGACATCGGCCACCC	644
Db	520	TATTACGAGTCMAAGGCCCAATTCATGTGCTGTCAAGCAAGAAATTCGGTCTCGG	579
QY	645	AAGACGGCGGCTCTTCCCTGCCCGGACCCAGCAATGCGCTGAGAGTGGGCGCGTGTG	704
Db	580	AAATCTGGGGCTCTTCTCCAGGTTGGGCTCTGTCTCGCTCCAGGACGAGACAGACAAG	639
QY	705	GCCCTTTCACGCCGAGAGCCGGGAGAACCGTGTCTGGCCATGGGGAGATGGGAGCCCC	764
Db	640	GCCCTGGAAGGACCTGAACCCCGGAGAAAGTGTCTGGCGACAGACGCCGGGAAACTCG	699
QY	765	ACCTTCACGATGTGCTCATTTTTCGTGAGACCGGAGACCCACAGGCTGAGACCTTTCAG	824
Db	700	GTTGTCAGCACTTCATCATGTTCACAGACCGAGACTCCACAGACGGAGTGTGTTTAC	759
QY	825	GTCATCGAGACTCAGGAGCCCCCAACGCCGCGCTGGCACTACACCGCGTCACTGCTTT	884
Db	760	GTCATGAGAAACGCAAGAACCCGTTGAAGAAATCACCTTACCGCGCTCACTCCTCTTT	819
QY	885	ACGGCTGACATCACACGAGACCGGACGCCGCTTCCGGGCCATTTTCCACGCCACGTG	944
Db	820	GTCCTCGCAACTCAACGAGGAAGATCTCCACACCATACCGCGCGGTATCCACGACGTGC	879
QY	945	CAGCCTGGCCAGTACGATGCGTGGTGGGCGTCCAGGGCTCG --- AGCCTGCCCGGCTG	1001
Db	880	AGAGCCGGACAAAAGGTGATGTGTTGATGATAGCGGTGAGTTAAATCTCTACCTG	939
QY	1002	GCACGTGCTCTACACAGTGGGCTCGGGGCTACAGCCCGCTCCAAAGAGATGGAGCA	1061
Db	940	CAGGGATATACAGGAGAGACAGCGGGGCTGTTGGCACAAGTAGTCAATGGAGAC	999
QY	1062	CTGGTGGTGAAGATGTGTGGCATCTCTCGGCGCGGTGGCTGACCAACCACTGGCT	1121
Db	1000	ATTGTGGTCGACAAATCTGGGCTCCTGTTAGCGCGTAATGAGAGACCAAGGGCTTGG	1058
QY	1122	CAGTTGGCTTCTGGCCCCCGAGACACTTTCACAGCTTGGCAT	1164
Db	1060	CATTGGCTTCTGGCCCCCGAGGCTTATTTATTCGTGCAT	1102

```
Search completed: June  5, 2000, 07:56:48
Job time: 3779 sec
```


CC constructs for altering hedgehog gene regulatory sequences) and
 CC ptc therapeutics (i.e., agents which mimic the effect of naturally
 CC occurring hedgehog proteins on patched signalling) that are
 CC effective in both human and animal subjects. A bioactive
 CC polypeptide comprising amino acid residues 28-202 of human Ihh is
 CC preferred. The products can also be used for the maintenance of
 CC differentiated neurons in cultures, and to enhance the implantation
 CC of such neuronal cells in an animal. They can also be used to
 CC prevent or treat neurodegenerative conditions arising from the use
 CC of certain drugs, and in the prevention and/or treatment of hypoxia,
 CC e.g., as a neuroprotective agent.

SO Sequence 1622 BP, 277 A, 549 C, 510 G, 286 T;

Query Match 100.0%; Score 1622; DB 1; Length 1622;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATCAGCCCAACGAGAGACCTCGCCCGCTCCCGGCTCCCGGCTCCCGATGCTCCCG 60
 DB 1 CATCAGCCCAACGAGAGACCTCGCCCGCTCCCGGCTCCCGGCTCCCGATGCTCCCG 60
 QY 61 CCGGGCTCCGGCCCGGCTCCCGGCTCCCGGCTCCCGGCTCCCGGCTCCCGGCTCCCG 120
 DB 61 CCGGGCTCCGGCCCGGCTCCCGGCTCCCGGCTCCCGGCTCCCGGCTCCCGGCTCCCG 120
 QY 121 CGGAGTGGGCTCGGGCCGCTCGGGTGGTGGGAGACCCCGGCGACCCGACGCAAC 180
 DB 121 CGGAGTGGGCTCGGGCCGCTCGGGTGGTGGGAGACCCCGGCGACCCGACGCAAC 180
 QY 181 TCGTGGCGCTCGGCTTACAGCAAGTTCAGCCCAATGTGCCGAGAGACCTTGGGCGCA 240
 DB 181 TCGTGGCGCTCGGCTTACAGCAAGTTCAGCCCAATGTGCCGAGAGACCTTGGGCGCA 240
 QY 241 GCGAGCGCTTGAAGGCAATGCTCGAGCTCCGAGCGCTTCAAGAGTCAACCCCA 300
 DB 241 GCGAGCGCTTGAAGGCAATGCTCGAGCTCCGAGCGCTTCAAGAGTCAACCCCA 300
 QY 301 ATTACATCCAGACATCATCTTCAAGAGAGAGAGAAACAGAGCGCCGCTCATGA 360
 DB 301 ATTACATCCAGACATCATCTTCAAGAGAGAGAGAAACAGAGCGCCGCTCATGA 360
 QY 361 CCGAGCGCTTGAAGAGAGCGCTTGAAGTCTGCTGCTATCTGCTATGAACAGAGCGCG 420
 DB 361 CCGAGCGCTTGAAGAGAGCGCTTGAAGTCTGCTGCTATCTGCTATGAACAGAGCGCG 420
 QY 421 GTGTGAAGCTCGCGGTACCGAGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
 DB 421 GTGTGAAGCTCGCGGTACCGAGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
 QY 481 TGCATTTGAGAGGCGCGGCTGGACATCACACATCAGACCGGAGCAATTAAGTATG 540
 DB 481 TGCATTTGAGAGGCGCGGCTGGACATCACACATCAGACCGGAGCAATTAAGTATG 540
 QY 541 GACTGCTGGCGGCTTGGAGTGGAGGCGCGCTTGAAGTGGGTATTAAGAGTCAAGG 600
 DB 541 GACTGCTGGCGGCTTGGAGTGGAGGCGCGCTTGAAGTGGGTATTAAGAGTCAAGG 600
 QY 601 CCGAGCGCTTGAAGTGGAGTGGAGGCGCGCTTGAAGTGGGTATTAAGAGTCAAGG 660
 DB 601 CCGAGCGCTTGAAGTGGAGTGGAGGCGCGCTTGAAGTGGGTATTAAGAGTCAAGG 660
 QY 661 TCCGTCGCGAGAGCGGCTTGAAGTGGAGGCGCGCTTGAAGTGGGTATTAAGAGTCAAGG 720
 DB 661 TCCGTCGCGAGAGCGGCTTGAAGTGGAGGCGCGCTTGAAGTGGGTATTAAGAGTCAAGG 720
 QY 721 GCGCGGAGAGCGGCTTGAAGTGGAGGCGCGCTTGAAGTGGGTATTAAGAGTCAAGG 780
 DB 721 GCGCGGAGAGCGGCTTGAAGTGGAGGCGCGCTTGAAGTGGGTATTAAGAGTCAAGG 780
 QY 781 TCATTTCTGAGAGCGGAGCGGCTTGAAGTGGAGGCGGCTTGAAGTGGGTATTAAGAGTCAAGG 840
 DB 781 TCATTTCTGAGAGCGGAGCGGCTTGAAGTGGAGGCGGCTTGAAGTGGGTATTAAGAGTCAAGG 840

QY 841 ACCCCACAGCGCGCTGGACATCAGACCCGCTGCTCTTTAGGCTGACAAATCACA 900
 DB 841 ACCCCACAGCGCGCTGGACATCAGACCCGCTGCTCTTTAGGCTGACAAATCACA 900
 QY 901 CCGAGCGGAGCGCGCTTCCGGGCGCAATTTGCGAGCCAGCGAGAGCTGGGCAATAG 960
 DB 901 CCGAGCGGAGCGCGCTTCCGGGCGCAATTTGCGAGCCAGCGAGAGCTGGGCAATAG 960
 QY 961 TGGTGGGCTGGGAGGCGAGGCTGAGCGCGCGCGGAGGAGGCTGCTACACAG 1020
 DB 961 TGGTGGGCTGGGAGGCGAGGCTGAGCGCGCGCGGAGGAGGCTGCTACACAG 1020
 QY 1021 TGGCCCTCGGGGCGCTACGCGCGCTACAAAGCATGGAGCATGTGTGAGAGATG 1080
 DB 1021 TGGCCCTCGGGGCGCTACGCGCGCTACAAAGCATGGAGCATGTGTGAGAGATG 1080
 QY 1081 TGGCATCTCTTCCGGGCGGCTGGTACACACCTGCTCAGTTGCTTGGCGCC 1140
 DB 1081 TGGCATCTCTTCCGGGCGGCTGGTACACACCTGCTCAGTTGCTTGGCGCC 1140
 QY 1141 TGAAGCTCTTACAGAGTGGCATGGGAGAGTGAACCCCGGGGAGGCTGCTATG 1200
 DB 1141 TGAAGCTCTTACAGAGTGGCATGGGAGAGTGAACCCCGGGGAGGCTGCTATG 1200
 QY 1201 ACCCCAGCTGCTTACCGCGCTGGGCGCTCTCTGTAGAGAGGCGAGCTTCCAC 1260
 DB 1201 ACCCCAGCTGCTTACCGCGCTGGGCGCTCTCTGTAGAGAGGCGAGCTTCCAC 1260
 QY 1261 TGGGATGTCCGGGCGAGGAGTGAAGAGTCAACCGCTGCTTGAAGTGTGT 1320
 DB 1261 TGGGATGTCCGGGCGAGGAGTGAAGAGTCAACCGCTGCTTGAAGTGTGT 1320
 QY 1321 ACTGGGTCAGAGAGCTTACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
 DB 1321 ACTGGGTCAGAGAGCTTACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
 QY 1381 GGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
 DB 1381 GGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
 QY 1441 AACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 DB 1441 AACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 QY 1501 AGGGAGTGTGTGAG 1560
 DB 1501 AGGGAGTGTGTGAG 1560
 QY 1561 CTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
 DB 1561 CTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
 QY 1621 CC 1622
 DB 1621 CC 1622

RESULT 3

X25104
 ID X25104 standard; cDNA; 1622 BP.
 AC X25104;
 DT 05-JUN-1999 (first entry)
 DE Human indian hedgehog protein Ihh cDNA.
 KW Indian hedgehog; Ihh gene; hedgehog therapeutic;
 KW ptc therapeutic; patched; signal transduction; muscle atrophy;
 KW carcharia; muscular myopathy; myoblastic sarcoma; therapy; ds.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT 51..1286
 FT CDS
 PN NC0910004-A2.
 PD 04-MAR-1999.

ID	Accession	Standard	CDNA	1281 BP
DT	DT	05-JUL-1999	(first entry)	
DE	Mouse indian hedgehog protein 11h cDNA.			
KW	Indian hedgehog; 11h gene; mouse; hedgehog therapeutic;			
KW	ptc therapeutic; patched; signal transduction; muscle atrophy;			
KW	cachexia; muscular myopathy; myoblastic sarcoma; therapy; ds.			
OS	Mus sp.			
FT	Key	Location/Qualifiers		
FT	CDS	1..1236		
FT		/*tag= a		
PN	PD	W09910004-A2.		
PD	PD	04-MAR-1999.		
PF	PD	28-AUG-1998: U17922.		
PR	29-AUG-1997:	US-057394.		
PA	(ONTO-)	ONTOGENE INC.		
PI	Bladgen CS, Currie PD, Hughes SM, Ingham PW;			
DR	WPI: 99-243557/20.			
DR	P-PSDB: Y05512.			
PT	A new method to regulate muscle growth			
PS	Disclosure; Page 96-98; 130pp: English.			
CC	This nucleotide sequence comprises a coding region for the mouse			
CC	Indian hedgehog protein 11h (see Y05512). The invention relates to			
CC	a method for modulating the formation and/or maintenance of muscle			
CC	tissue by ectopically contacting muscle cells, especially muscle			
CC	stem/progenitor cells, in vitro or in vivo, with a hedgehog			
CC	therapeutic (i.e. hedgehog polypeptides and gene therapy			
CC	constructs) or ptc therapeutic (i.e. a small organic molecule that			
CC	mimics the effect of hedgehog proteins on patched signalling, or			
CC	activates or potentiates patched signalling) in an amount effective			
CC	to alter the growth state of the treated cells. Also claimed is a			
CC	method for treatment or prevention of disorders of, or surgical or			
CC	cosmetic repair of, such muscle tissues, by administering a			
CC	hedgehog polypeptide or ptc therapeutic. The disorder may be			
CC	muscle atrophy, in particular skeletal muscle atrophy or cardiac			
CC	muscle atrophy, cachexia, or muscular myopathy (all claimed). The			
CC	hedgehog polypeptide or ptc therapeutic can inhibit growth of			
CC	myoblastic-derived tissue to provide treatment of hyperplastic or			
CC	neoplastic growth of muscle tissue such as in myoblastic sarcoma			
CC	(also claimed). The hedgehog therapeutic preferably comprises at			
CC	least a bioactive extracellular portion of a hedgehog protein (see			
CC	Y05510-19) encoded by a vertebrate hedgehog gene (see X25098-107),			
CC	especially a human hedgehog gene.			
SQ	Sequence	1281 BP; 229 A; 401 C; 393 G; 258 T;		
Query Match	63.5%; Score 1029.8; DB 1; Length 1281;			
Best Local Similarity	88.4%; Pred. No. 2.9e-194;			
Matches 1129; Conservative	0; Mismatches 147; Indels 1; Gaps			
0Y	51 ATGTCTCCGCGCGGCTCCGCGGCCCGGAGCTGCATCTTGCTGTGCTGTGCTGCTG	110		
DB	1 AATGTTCTCCGCTGAGCTCCGCGGCCCGGAGTCTGTGCTGTGCTGTGCTGCTGCTCT	60		
0Y	111 GTGTGTCCCGCGGCGATGGGGGCTGGGGGCGGGGTGGGGTGGTGGGGACCCCGCCGACCG	170		
DB	61 GTGTGTCCGGGGGCGGGGCGGGGCGGGGCGGGGCGGGGTGGGGACCCCGCGAGCCG	120		
0Y	171 CCAGCCAAATCTGTCCGCGCTGCCTACAGAGAGTTTCAGGCCCAATGTGCCGAGAAAGC	230		
DB	121 CCTGCGAAGCTGTGCTCTTCTGCTTACAGAGATTTACGCCCAAGTGGCCGAGAAAGCC	180		
0Y	231 CTGGGCGCGCAGCGAGCTATGAAGAGCAAGATCGCTGCGACCTCGAGCGCTTCAAGAG	290		
DB	181 CTGGGCGCGCAGGGGCGGTAGAGAGGCAAGATCGGGGCGACCTGTGAGCGTTCAAGAG	240		
0Y	291 CTGACCCCAATTACATTCAGACATCATCTTTCAAGAGACAGAGAAACACAGCGCCGAC	350		
DB	241 CTCACCCCAATCAATCAATCCGACATCATCTTCAAGAGACGAGGAACAGCGGTGCCGAC	300		
0Y	351 GCGCTCATGAGCCACCGCTGGAAGAGCGGCTGAACCTGCTGGTATCTGGGTATGAG	410		
DB	301 GCGCTCATGAGCCACCGCTGGAAGAGCGGCTGAACCTGCTGGTATCTGGGTATGAG	360		

QY	411	CAGTGGCCCGGTGTGAAGCTGCGCGGGTGGACGAGAGGGCTGGGACGAGAGACCGCCACATCTCA	470
Db	351	CAGTGGCCCTGTGTGAAACACTGCGGGTGTACCCGAGAGCGCGGGATGAAGATGGCATATCTACTCA	420
QY	471	GAGGAGTCCCTGCATTATATGAGGGCCGCGGGGTGGACATCACCATCAGACCGCAGCCG	530
Db	421	GAGGAGCTTTTACACTATGAGGGCCCGCGGGTGGATATCACCATCTCAGACCGTACCGA	480
QY	531	AATAGATGAGACTGCGTGGCGGCTTGGACATGGAGAGCCGCTTGACTGGGTGATATAC	590
Db	481	AATAGATGAGACTGCGTGGCGGCTTGGACATGGAGAGCCGCTTGACTGGGTGATATAC	540
QY	591	GAGTCAAGGGCCCACTGCACTTGTCTCTGTCAGTCCGACACTCTGCGCGCAGCCAAAGAG	650
Db	541	GAGTCAAGGGCCCACTGCACTTGTCTCTGTCAGTCCGACACTCTGCGCGCAGCCAAAGAG	600
QY	651	GCGCGCTGCTTCCCTCGCGGAGCCCAAGTACGCTTGAGAGTGGGGCGCGCTGTGCTTG	710
Db	601	GGTGGCTGCTTTCCTGCGCGAGCCCAAGTGGCGCTTGAGAGAGGGGAGAGTGTGGCCCTG	660
QY	711	TCACCGCGAGGGCCGGGAGACGCTGTGTGGCCATGGGGGAGAGATGGAGAGCCCAACTTC	770
Db	661	TCACGTGTAAAGCCAGGAGAGCCGGGTGTGGCCATGGGGGAGAGATGGAGAGCCCAACTTC	720
QY	771	AGCGATGTGCTATTTTTCCGTGAGACCCGAGCCCAAGGCTAGAGACCTTCCAGGTATAC	830
Db	721	AGTATGTGCTATTTTTCCGTGAGACCCGAGCCCAAGGCTAGAGACCTTCCAGGTATAC	780
QY	831	GAGACTAGAGACCCCGCAGCCGCTTGGCACTACACCCGCTACACTGCTTTAGGCT	890
Db	781	GAGACTAGAGACTCTCCGCGTGGGTGGCTACACCGCTCCCACTGCTTTGATTTGCG	840
QY	891	GACATACACAGGAGCCGGAGGCGGCTTCCGGGCCACATTGGCAGCCAGCTGAGAGCT	950
Db	841	GACATACATTAAGACAGCAGGACCCCACTTCCGGGCCCAATTGGCAGCCAGCTGAGACCA	900
QY	951	GGCCAGTACGCTGTGTGGGTGGGTGCCAGGAGCTTCACACTGCGCGCTGGCAGACTGTC	1010
Db	901	GGCCAAATATGTGCTGATCAGGGGTATCAAGAGCTTCACACTGCTGGGTGGCAGACTGTC	960
QY	1011	TCCTACACAGCTGGCCCTCGGGGGCTTCAGCCCGGCTCACAAAGCATGGGACACTGCTGTG	1070
Db	961	TCACACCCAGCGGGCCCTTGGGTGCTATGCTCTCCACAAAGCATGGGACACTTGTGTG	1020
QY	1071	GAGATGTGGGTGATCCTCTTCGGGGCGTGGGCTGACACACACTGGCTAGTTGGCC	1130
Db	1021	GAGATGTGGGTGCTCTCTCTTTCGACGTGTGTGGACACACACTGTGGCTAGTTGGCC	1080
QY	1131	TTTCGGCCCTTGAGACTTTTTCACAGCTTGGCATGGGGCAGCTGGACCCCGGGAGGGGT	1190
Db	1081	TTTCGGCCCTTGAGACTTTTTCACAGCTTGGCATGGGGCAGCTGGACCCCAAGTAAAGGT	1140
QY	1191	GTCGATGTGTAACCCCAAGCTGCTCTACCGGCTTGGGGCTCTCTGCTGTAAGAGGGGAC	1250
Db	1141	GTCGATCTCCCACTCCAGATGCTCTACCGGCTTGGGGCTCTCTGCTGTAAGAGGGGAC	1200
QY	1251	TTTCACCACTGGGGCAGTCCGGGGGAGGGAGCTGAAGAGATC-CACCGCTGGCTCTCT	1309
Db	1201	TTTCATTCACCTGGGGCAGTGTGGGGGAGGGAGCTGAAGAGACTTACACACTTGGCTCTCT	1266
QY	1310	GGAACTGCTACTACTGGG 1326	
Db	1261	GGAAGTGTGTGGCTGG 1277	
RESULT	7		
Q91641			
ID	Q91641	standard; cDNA; 939 BP.	
AC	Q91641		
DEF	18-MAR-1996	(first entry)	
DE	Human Indian hedgehog protein gene.		
KW	Human; Indian hedgehog protein; nested polymerase chain reaction; PCR.		

KW fetal lung; probe; primer; diagnostic; nervous system disorder;
 KM gene therapy; antibody; ss.
 OS Homo sapiens.
 PN W09518856-A1.
 PD 13-JUL-1995.
 PF 30-DEC-1994; U14992.
 PR 30-DEC-1993; US-176427.
 PR 14-DEC-1994; US-356060.
 PA (HARD) HARVARD COLLEGE.
 PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
 PI Ingham PW, McMahon AP, Tabin CJ;
 DR WPI: 95-255060/33.
 DR P-PSDB: R77344.
 PT Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful
 PT to treat degenerative nervous system disorder(s) and in gene
 PT therapy.
 PS Claim 4; Page 146-47; 210pp; English.
 CC The sequence encodes a human Indian hedgehog protein, homologous
 CC to a Drosophila hedgehog protein (R77337), and has been isolated by
 CC screening of human genome DNA by nested polymerase chain reaction
 CC using primers 091643, 091644 and 091645, followed by use of a clone
 CC lambda gt-10. Probes and primers derived from hedgehog sequences
 CC may be used as diagnostic agents for neuromuscular, autonomic or
 CC central nervous system disorders, and the gene may also be used in
 CC gene therapy. Antibodies generated from the encoded protein may be
 CC used as therapeutic or research reagents.
 SQ Sequence 939 BP; 154 A; 305 C; 314 G; 166 T;

Query Match 57.7%; Score 936; DB 1; Length 939;
 Best Local Similarity 100.0%; Pred. No. 7.6e-176;
 Matches 936; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 351 GCGCTCATGACCCAGCGCTGCAAGAGCCGCTGAAGTCTGCTGCTATCTCGGTATGAC 410
 DB 4 CGCCCATGACCCAGCGCTGCAAGAGCCGCTGAAGTCTGCTGCTATCTCGGTATGAC 63

QY 411 CAGTGGCCCGGTGTGAGCTGCGGGGTGACCGAGGCTGGAGAGGAGCGGACCACTCA 470
 DB 64 CAGTGGCCCGGTGTGAGCTGCGGGGTGACCGAGGCTGGAGAGGAGCGGACCACTCA 123

QY 471 GAGGAGTCCCTGCTATGAGAGCGCGCGGTGAGCATCACCACATCAGACCGGACCGC 530
 DB 124 GAGGAGTCCCTGCTATGAGAGCGCGCGGTGAGCATCACCACATCAGACCGGACCGC 183

QY 531 AATAAGTGTGAGCTGCTGCGCGCTTGGCACTGAGAGCGCGCTTGGCTGATTAAC 590
 DB 184 AATAAGTGTGAGCTGCTGCGCGCTTGGCACTGAGAGCGCGCTTGGCTGATTAAC 243

QY 591 GAGTCAAGGCGCCAGCTGCTGCTGCTCAAGTCCGAGCACTGCGGCGGACCAAGCG 650
 DB 244 GAGTCAAGGCGCCAGCTGCTGCTGCTCAAGTCCGAGCACTGCGGCGGACCAAGCG 303

QY 651 GCGCGCTCTTCCCTGCGGAGCCAGGTAGCGCTGAGAGTGGGGCCGCTGTGGCCTTG 710
 DB 304 GCGCGCTCTTCCCTGCGGAGCCAGGTAGCGCTGAGAGTGGGGCCGCTGTGGCCTTG 363

QY 711 TCAGCGGTGAGGCGGAGAGCCGTGCTGCTGCGCATGSGGAGAGTGGAGCCCACTTC 770
 DB 364 TCAGCGGTGAGGCGGAGAGCCGTGCTGCTGCGCATGSGGAGAGTGGAGCCCACTTC 423

QY 771 AGCGATGCTCATTTCTTCTGAGCGGAGCGCCAGCGTGAAGAGCTTCCAGGTATC 830
 DB 424 AGCGATGCTCATTTCTTCTGAGCGGAGCGCCAGCGTGAAGAGCTTCCAGGTATC 483

QY 831 GAGACTCAGAGCCGCCAGCGCGCTGCGCATCAGACCCGCTCAGCTCTTTACGGCT 890
 DB 484 GAGACTCAGAGCCGCCAGCGCGCTGCGCATCAGACCCGCTCAGCTCTTTACGGCT 543

QY 891 GACATATCAGAGCGGAGCGGAGCCGCTTCCGGGCGCATTTGCCAGCAGTGCAGCT 950
 DB 544 GACATATCAGAGCGGAGCGGAGCCGCTTCCGGGCGCATTTGCCAGCAGTGCAGCT 603

QY 951 GCGCAGTACGAGCTGATGCTGGGGGTGCGCAGGCGCTGAGCGCGCGTGGAGCTGTC 1010
 DB 604 GCGCAGTACGAGCTGATGCTGGGGGTGCGCAGGCGCTGAGCGCGCGTGGAGCTGTC 663

QY 1011 TCTACACAGCTGGGCGCTGAGGCGCTTACGCGCGCTCAAGAGCATGGAGACTGTGTG 1070
 DB 664 TCTACACAGCTGGGCGCTGAGGCGCTTACGCGCGCTCAAGAGCATGGAGACTGTGTG 723

QY 1071 GAGATATGCTGATCTGCTGCTGCGCGCGCTGAGCTACCAACACTGCTCACTTGGCC 1130
 DB 724 GAGATATGCTGATCTGCTGCTGCGCGCGCTGAGCTACCAACACTGCTCACTTGGCC 783

QY 1131 TTCTGCGCCCTGAGACCTCTTACAGCTTGGCATGAGGAGGAGCGGAGAGGCT 1190
 DB 784 TTCTGCGCCCTGAGACCTCTTACAGCTTGGCATGAGGAGGAGCGGAGAGGCT 843

QY 1191 GTGCTATGTAACCCAGCTGCTTACCGCGCTGGGCGCTCTCTGTAGAAGAGCGAGC 1250
 DB 844 GTGCTATGTAACCCAGCTGCTTACCGCGCTGGGCGCTCTCTGTAGAAGAGCGAGC 903

QY 1251 TTCCACCCACTGGGCGCATGTCGGGCGCAGGAGAGCTGA 1286
 DB 904 TTCCACCCACTGGGCGCATGTCGGGCGCAGGAGAGCTGA 939

RESULT 8
 091640
 ID 091640 standard; cDNA; 1056 BP.
 AC 091640;
 DT 14-MAR-1996 (first entry)
 DE Mouse Indian hedgehog protein gene.
 KM Mouse; Indian hedgehog gene; probe; primer; diagnostic;
 OS nervous system disorder; gene therapy; antibody; ds.
 PS Mus musculus.
 FH Key Location/Qualifiers
 FT cds 1..1008
 FT /product= Mouse Indian hedgehog protein
 PN W09518856-A1.
 PD 13-JUL-1995.
 PF 30-DEC-1994; U14992.
 PR 30-DEC-1993; US-176427.
 PR 14-DEC-1994; US-356060.
 PA (HARD) HARVARD COLLEGE.
 PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
 PI Ingham PW, McMahon AP, Tabin CJ;
 DR WPI: 95-255060/33.
 DR P-PSDB: R77343.
 PT Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful
 PT to treat degenerative nervous system disorder(s) and in gene
 PT therapy.
 PS Claim 4; Page 137-38; 210pp; English.
 CC The sequence encodes a mouse Indian hedgehog protein, homologous
 CC to a Drosophila hedgehog protein (R77337), and has been isolated by
 CC low stringency screening of a mouse genome DNA library and
 CC screening of an 8.5 day post coitum cDNA library. The partial cDNA
 CC is complete at the 3'-end, as evidenced by the presence of a
 CC polyadenylation consensus sequence and short poly-A tail. Probes
 CC and primers derived from hedgehog sequences may be used as
 CC diagnostic agents for neuromuscular, autonomic or central nervous
 CC system disorders, and the gene may also be used in gene therapy.
 CC Antibodies generated from the encoded protein may be used as
 CC therapeutic or research reagents.
 SQ Sequence 1056 BP; 204 A; 321 C; 312 G; 219 T;

Query Match 52.0%; Score 843.2; DB 1; Length 1056;
 Best Local Similarity 88.2%; Pred. No. 1.3e-157;
 Matches 928; Conservative 0; Mismatches 123; Indels 1; Gaps 1;

QY 276 GAGGCTTCAAGAGCTACCCCAATTAATCAGATCATCTTCAAGAGAGAGAG 335
 ||||||||| ||||||||| ||||||||| ||||||||| |||||||||

```

Db      1  GAGCGCTTCAAGAGCTCACCCCACTACATCCGACATCATCTTCAAGAGAGAG 60
QY      336 AACACAGAGCCGACCCGCTCATGACCCGCTGCAAGAGACCGGCTGCTGGCT 395
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      61 AACACAGAGGTCGCCGACCCGCTCATGACCCGCTGCAAGAGACCGCTGCTGACTCTG 120
QY      396 ATCTCGGTGATGAACAGTGGCCGCTGTGAAGCTGGGGGTGACCGAGGGCTGGGAG 455
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      121 ATCTCTGTATGAACAGTGGCCGCTGTGAAGCTGGGGGTGACCGAGGGCTGGGAG 180
QY      456 GAGGCGCCACACTCAGAGAGTCCCTCATATATAGAGGCGCGGGGTGAGACATACACA 515
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      181 GAGGCGCCATACTCAGAGAGTCTTTACATATAGAGGCGCGGGGTGATATACACACC 240
QY      516 TCAGACCGGACCGCATATAGTATGAGTGTGGCGGCTTGGAGTGGAGGCGGGCTT 575
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      241 TCAGACCGTACCGAATATATATAGTGTGGCGGCTTGGAGTGGAGGCGGGCTT 300
QY      576 GACTGGGTATTATAGATGCAAAAGGCGGAGTGCATGTCTCCGTCAGTGCAGACTCG 635
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      301 GACTGGGTATTATAGATGCAAAAGGCGGAGTGCATGTCTCCGTCAGTGCAGACTCG 360
QY      636 GCGGCAACCAAGAGCGGCGGCTTCCCTGCGGAGCCCAAGTACCCCTGGAGATGG 695
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      361 GCGGCTGCGCAAGAGTGGCTGTCTTCTGCGGAGCCCGAGTGCCTAGAGAACGGG 420
QY      696 GCGGCTGCGCTTGTAGGCGGTGAGGCGGAGACCGTGTGCTGGGCGATGGGAGGAT 755
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      421 GAGGCTTGGCGCTGTAGGCTGTAAAGCCAGAGAACCGGCTGTGGCGATGGGAGGAT 480
QY      756 GGAAGCCCACTTACAGATGTGCTATTTCTTGAGACCGGAGCCCAAGAGCTAGA 815
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      481 GGAAGCCCACTTACAGATGTGCTATTTCTTGAGACCGGAGCCCAAGAGCTAGA 540
QY      816 GCCTTCAGATCATGAGACTCAGAGACCCCGGAGCGGCTGGAGCTCAGCCGCTAC 875
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      541 GCTTCCAGGTCATGAGACTCAGAGATCTCCGCTGGGCTGCGCTCAGCCGCTGAC 600
QY      876 CTGCTCTTACGCGTGCATACAGAGAGCGGAGCGGCGCTTCCGCGGACATTTGGC 935
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      601 CTGCTCTTACGCGTGCATACAGAGAGCGGAGCGGCGCTTCCGCGGACATTTGGC 660
QY      936 AGCCACGTGAGCGCTGAGCAGTACGCTGTGGCTGGGCTGAGGCGCTGAGCGCTGCC 995
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      661 AGCCATGTGACAGCGGCAATATGTGTGTATCAGGGGTATCAGAGGCTTCCAGGCT 720
QY      996 CGGCTGAGCTGTCTTACACAGCTGGGCTCGGGGCTTACGCCCGGCTCAGAAAGCAT 1055
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      721 CGGCTGAGCTGTCTTACACAGCTGGGCTCGGGGCTTACGCCCGGCTCAGAAAGCAT 780
QY      1056 GGAAGCATGTGTGTGAGAGATGTGTGATCCGCTGCTCGGGCGGCTGTGAGCAGCAC 1115
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      781 GGAAGCATGTGTGTGAGAGATGTGTGATCCGCTGCTCGGGCGGCTGTGAGCAGCAC 840
QY      1116 CTGCTCAGTGTGGCTTGTGGCGCTGAGACTTTTCAACAGCTTGGCATGGGAGCTGG 1175
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      841 CTGCTCAGTGTGGCTTGTGGCGCTGAGACTTTTCCCAATTTGGCATGGGAGCTGG 900
QY      1176 ACCCGGGGAGGTGTGATGTGATCCCGGAGCTGCTTACGCCCTGGGCGTCTCGT 1235
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      901 ACCCGAGTGTGAGGTGTGATGTGATCCCGGAGCTGCTTACGCCCTGGGCGTCTCGT 960
QY      1236 CTAAAGAGGAGCTTCCACACCACTGGGAGTGTCCGGGGAGAGGAGCTGAAGAGCTC- 1295
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      961 CTAAAGAGGAGCTTCCACACCACTGGGAGTGTGTGGGAGAGAGGAGCTGAAGAGCTC 1020
QY      1295 CACGCTGCTCTGAGAACTGTGTACTGGG 1326
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1021 AACCACTGCTCTCTGAACTGTGTGTG 1052

```

```

ID      091636 standard; cDNA; 1277 BP.
AC      091636;
DT      22-FEB-1996 (first entry)
DE      Chicken sonic hedgehog protein gene.
KW      Chicken; sonic hedgehog gene; nested polymerase chain reaction;
KW      stage 22/22 limb bud; transgenic mouse screening; probe; primer;
KW      PCR; diagnostic; nervous system disorder; gene therapy; antibody;
KW      ds.
OS      Gallus domesticus.
PN      W09518856-A1.
PD      13-JUL-1995.
PF      30-DEC-1994; U14992.
PR      30-DEC-1993; US-176427.
PR      14-DEC-1994; US-356060.
PA      (HARD ) HARVARD COLLEGE.
PA      (TMCR ) IMPERIAL CANCER RES TECHNOLOGY.
PI      Ingham PW, McMahon AP, Tabin CJ;
DR      WPI; 95-255060/33.
DR      P-PSDB; R77338.
PT      Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful
PT      to treat degenerative nervous system disorder(s) and in gene
PT      therapy.
PS      Claim 2; Page 133-35; 210pp; English.
CC      The sequence encodes a chicken sonic hedgehog protein, homologous
CC      to a Drosophila hedgehog protein (R77337), and is isolated by
CC      nested polymerase chain reaction using primers VH50 (091643),
CC      VH30 (091644) and VH31 (091645). A clone resulting from the
CC      amplification (encoding R77348) is used to isolate the full-length
CC      cDNA from a stage 22/22 limb bud cDNA library. Primer 091649 is
CC      derived from the chicken cDNA sequence, and is used to screen
CC      transgenic mice. Probes and primers derived from the sonic
CC      hedgehog sequence may be used as diagnostic agents for
CC      neuromuscular, autonomic or central nervous system disorders, and
CC      the gene may also be used in gene therapy. Antibodies generated
CC      from the encoded protein may be used as therapeutic or research
CC      reagents.
SQ      Sequence 1277 BP; 262 A; 396 C; 387 G; 232 T;

Query Match      28.3%; Score 459; DB 1; Length 1277;
Best Local Similarity 65.2%; Pred. No. 3; 9e-82;
Matches 763; Conservative 0; Mismatches 360; Indels 48; Gaps 4;

QY      163 GGCACCGCCAGCAAACTGCTGCGCTGCTGCAAGAGCTTCAAGAGCTTGGCG 222
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      107 GGAAGCAACCCCAAAAGCTGACCCCGCTTAAAGCAATTTATCCCAATGTGGCAG 166
QY      223 AGAAGACCTTGGGCGCCAGGAGGAGCTATGAAAGGCAAGATCGTCGAGCTCCGAGCGCT 282
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      167 AGAAGACCTTGGGCGCCAGGAGGAGTATGAAAGGAGATCACAAAGAACTCCGAGAGAT 226
QY      283 TCAAGAGCTCACCCCAATTACATCCAGACATCATCTTCAAGAGAGAGAGCAACAG 342
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      227 TTAAGAACTAACCCCAATTATCAACCTGATATTTTAAAGATGAAGAGAAACAGG 286
QY      343 GCGCGACCGGCTCATATACCAAGCGCTGCAAGAGCGGCTGAACTCGCTGCTATTCGG 402
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      287 GAGCTGACAGACTATGATCTCAGCGCTGCAAGAGCAAGCTGAATGCTTGGCATTTCCG 346
QY      403 TGATGAACCACTGGCCCGGTGTGAAGCTGGGAGTGGACGAGAGGCTGGAGAGAGAGCGCC 462
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      347 TGATGAACCACTGGCCCGGTGTGAAGCTGGGAGTGGACGAGAGGCTGGAGAGAGATGGCC 406
QY      463 ACCACTCAGAGAGTCCCTGATATATAGAGGCGCGCGGTGTGATATACACATCAGAGCC 522
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      407 ATCACTCCAGAGAACTCCTGACTACAGAGGCTCCGCGCTGTGACATACCACTCGGATC 466
QY      523 GCGACCGCAATATATGAGTGTGCTGGCGGCTTGGAGTGGAGAGGAGCGGCTTGAAGTGG 582
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      467 GGAACCGAGCAATGAGAGTGTGGCCCGCTTCGCGCTGAGAGGCGGCTTCACTGTGG 526
QY      583 TGTATTAGAGTCAAGAGCCACGTCGATGCTCTCCGTAATCCGAGCACTGGCGCAG 642
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



```

Db 527 TCTACTACGATCCAGGCGCACATCCACTGCTCCGTCAGAGCAAGAACTGACGTCGAG 586
Qy 643 CCAAGAGCGGGGCTGCTCTCCCTGCGGAGAGCCAGCTAGCGCTGAGAGTGGGGCGCTG 702
Db 587 CGAATATCAGAGAGGCTGCTCTCCCTGCTGACGCCACAGTGCAGCTGGAGCATGGAGGACACA 646
Qy 703 TGGCCTGTGACCGGTGAGCGGGGAGACCGTGTGCTGCTGACATGGGGAGATGGAGCC 762
Db 647 ACCTGTGAAGAGACCTGAGCCCTGGGGAGCCGGGTGCTGCTGCTGACCGGAGCGCGGCG 706
Qy 763 CCACCTGACGATGTCATTTCTTCCTGAGCCGAGCCGACAGGCTGAGAGGCTTCC 822
Db 707 TCTCTACAGTATCTCTCTACCTCTCTCTGACCGGATGAGACAGCTCCGAAAGCTCTTCT 766
Qy 823 AGGTCTACAGATCAGAGACCCCGACCGCGCTGAGCATCACACCGCTGACCTGCTCT 882
Db 767 AGGTATGAGAGCGGGAGCCCGCGGGCTGCTGCTGAGAGCGGGGCGGACCTGCTCT 826
Qy 883 TTACGGCTGACATCAGACGAGCGG-----CAGCCGCTTCCGGGCGCAT 930
Db 827 TTGTGGCCCCCAGCAGCAACAGTCCGAGCGCACAGGCTCCACAGTGGCCAGGCGCTCT 886
Qy 931 TTGGCAGCAGCGAGCGCTGGGCA-----GTACGTGCTGGGCTGGGGTGGCCAGGCGC 984
Db 887 TCGCCAGCAGCGAGCGCTGGGCAAGCTGTCTATGTGCTGGGCGGAGCGGCGGACGACG 946
Qy 985 TCGAGCTGCGCGCGCTGAGCTGTCTCTACAC---ACGTGGCGCTTCCGGGCGCTAGCGCC 1041
Db 947 TCGTGGCGCGCTGTCTCTACACAGGCTGTCTATGTGGGAGAGAGGCTGCGAGGCTAGCGCC 1006
Qy 1042 CGCTCAGAAAGCATGGGACACTGTGTGTGAGAGATGTGGTGGCATCTCTTCCGGCGCG 1101
Db 1007 CACTCACCCCGCAGGCGACCATCTCTCATACCGGGGTGTGGCTCTCTCTACAGCTGTA 1066
Qy 1102 TGGCTACACCAACCACTGGGCTAGTTGGCTTGGCCCTTGGAGACTTTCACAGCTTGG 1161
Db 1067 TCGAGAGACAGTGTGGGCGCATTTGGGCGCTTGGCACCATTCCGCTTGGCTCAGGGGCTGC 1126
Qy 1162 CATGGGCGACGTGAGACCCCGGGGAG-----GGTGTGC 1194
Db 1127 TGGCGCGCTCTGCGCCAGATGGGGGCGCATGCCATCTGCGGACACACACACAGCTGGCATCC 1186
Qy 1195 ATTGGTACCCCAAGCTGCTCTACCGCTTGGGCGCTCTCTGCTAGAAAGGCGAGCTTCC 1254
Db 1187 ATTGGTACCTCAGGCGCTCTCTACCGCATGCGAGCTGGGTGCTGTGAGCGGCTGC 1246
Qy 1255 ACCCACTGGGCACTGTCCGGGCGAGGAGCTG 1285
Db 1247 ATCCGCTGGGCACTGTGCGACCGGCGCAGCTG 1277

RESULT 10
X16182
ID X16182 standard; cDNA: 1277 BP.
AC X16182;
DE 29-APR-1999 (first entry)
BT Chicken Shh hedgehog cDNA sequence.
DI patched; hedgehog; plc therapeutic; neuroprotective; neuronal cell;
KW brain infarction; cerebral infarction; transient ischemic attack;
KW stroke; cerebral infarct volume; spinal cord; oedema; trauma;
KW hemorrhage; encephalomyelitis; coronary bypass; cerebral hypoxia; ds.
OS Gallus sp.
FH key
FT 1. 1275
FT /tag= a
FT /product= "hedgehog sequence"
FT /note= "no stop codon given"
PN WO9900117-A2.
PD 07-JAN-1999.
PF 26-JUN-1998: U13387.
PI 27-JUN-1997: US-883656.
PA (ONTO-) ONTOGENY INC.
PI Mahanahappa NK.

```

```

DR WPI: 99-095458/08.
DR P-PSDB: W94468.
PT Method for limiting damage to neurons caused by ischemic or epoxic
PT conditions - is used for the treatment and prevention of e.g.
PT cerebral infarction, stroke and transient ischemic attacks
PT Disclosure: Page 63-64, 104pp. English.
PS A method has been developed for limiting the damage to neuronal cells by
PS ischemic or epoxic conditions by administering a plc (patched)
PS therapeutic agent to reduce cerebral infarct volume (CIV). Damage to
PS neuronal cells can also be limited by administering a gene activation
PS construct which recombines with the genomic hedgehog gene to provide a
PS heterologous transcription regulator linked to the coding region of this
PS gene. Administration of the plc therapeutic agent is used to protect
PS cerebral tissues against ischemic injury; to treat cerebral infarct or
PS ischemia, stroke (thrombotic or embolic) and transient ischemic
PS attacks. It may also be used as a prophylactic in many other cases of
PS injury to the brain or spinal cord, oedema caused by trauma, hemorrhage
PS and encephalomyelitis, or in conjunction with (coronary bypass) surgery.
PS treatment (which may be prophylactic) is used where ischemic/epoxic
PS conditions may cause cerebral hypoxia, or progressive loss of neurons
PS due to oxygen depletion, including in patients with hypotension. The
PS treatment reduces CIV by at least 25, particularly at least 70%. The
PS present sequence encodes a hedgehog sequence given in the present
CC invention.
CC Sequence 1277 BP: 262 A: 396 C: 387 G: 232 T:
SQ

Query Match 28.3%; Score 459; DB 1: Length 1277;
Best Local Similarity 65.2%; Pred. No. 3.9e-83;
Matches 763; Conservative 0; Mismatches 360; Indels 48; Gaps 4;

Qy 163 GCGAGCGCCAGCGAACTGTCGCTGCGCTGCTCAGACAGTTCAGCCCAATGTGCCG 222
Db 107 GGAGGACACCCAAAAGCTGACCGCTTACCTTAACAGATTATTCCTCAATGTGGCAG 166
Qy 223 AGAAGACCTGGGCGCCAGCGGAGCGCTTGAAGCAAGATGCTGCGACGCTCCAGCGCT 282
Db 167 AGAAGACCTTGGGCGCGGAGATGAAGAGAAATATCAAGAAATCTCCGAGAT 226
Qy 283 TCAAGAGCTACACCCCAATTACATCCAGACATCATCTTCAGAGAGAGAAACAG 342
Db 227 TTAAAGACTAACCCCAATTACAACTGACATTAATTTTAAGATTAAGAAACAG 286
Qy 343 GCGCGAGCGCTCATGACCCAGCGCTGCAAGAGACCGCTGAATCTGCTGCTATCTCG 402
Db 287 GAGCTGACAGACTGATGACTGACGCTGCAAGAGACCAATGATCCCTGCGATCTCG 346
Qy 403 TGATTAACAGTGGCGCGCTGTGAAGCTGGCGGTGACCGAGGCTGGAGACAGAGCGCC 462
Db 347 TGATTAACAGTGGCGCGCTGTGAAGCTGGCGGTGACCGAGGCTGGAGAGATGGCC 406
Qy 463 ACCACTAGAGAGTCCCTGATTTAGAGGCGCGCGGTGAGATCCACCATAGACC 522
Db 407 ATCACTCCGAGAGATCGCTGACATCGAGGGTCCGCGGTGAGATCCACAGTGGATC 466
Qy 523 GCGAGCGGAGTATGATGACTGTGCGCGCTTGGCAGTGGAGGCGGCTTGTACTGG 582
Db 467 GGGAGCGGAGAGTAAAGAAATGCTGGCGCGCTGCGCTGAGAGCGCGCTTGCAGTGG 526
Qy 583 TGTATTAGAGTCAAAAGGCCACAGTGTGCTGCTGCAAGTCCAGAGACCGCGCGAG 642
Db 527 TCTACTACGAGTCCAAAGCGCACATCCCTCCGTAAAGACAAATACTGATGGGAG 586
Qy 643 CCAAGAGCGGCGCTGCTCTCCCTGCGGAGCCAGAGTACGCTGAGAGTGGGGCGCTG 702
Db 587 CGAATATCAGAGAGGCTGCTCTCCCTGCTGACGCCACAGTGCAGCTGAGAGATGAGGACACA 646
Qy 703 TGGCCTGTGACCGGTGAGCGGGGAGACCGTGTGCTGCTGACATGGGGAGATGGAGACC 762
Db 647 AGCTGTGAAGAGCTGAGCCCTGGGAGACCGGTGCTGCTGCTGACGCGGAGCGCGCG 706
Qy 763 CCACCTGACGATGTCATTTCTTCCTGAGCCGAGCCGACAGGCTGAGAGCTTCC 822

```


Db 887 TCGCCAGCACTGAAAGCTGGCCAACTGTCTATGTCTGGGCCAGGGCGGCACTAC 946
 Qy 985 TCGAGCTCCCGCGTGGAGCTGTCTCTACAC---ACGTGGCCCTCGGGGCTTACGCC 1041
 Db 947 TCGTGGCGGCGTCTGTCCACAGAGGTCTCATGTGGGAGAGAGGGCTCGAGGCTACGCC 1006
 Qy 1042 CGCTCAAAAGATGGGACACTGTGTGTGAGAGATGTGTGATCTCTCTCGGGCCG 1101
 Db 1007 CACTCACCCCGGAGGACCATCTCATACACGGGTGTGGCTCTCTCTCTCTCTCTA 1066
 Qy 1102 TGGTACACCACTGAGTGTGAGTGTGGCCCTTGAGACTCTTCAAGCTTGG 1161
 Db 1067 TCGAGAGACACAGTGTGGCCCATGTGGCCCTTGACACATCTCGCTTGGCTCAAGGGCTGC 1126
 Qy 1162 CATGGGCGACGTGAGCCCGGGGGAG-----GGTGTGC 1194
 Db 1127 TGGCGCCCTCTGCCAGATGGGGGCATCCCTACTGCGCCACACACACACACTGGATCC 1186
 Qy 1195 ATTGTACCCCAAGCTGTCTACCGCTGGGCGTCTCTCTCTAAGAGGCGACTTCC 1254
 Db 1187 ATTGTACACAGGGCTCTCTACCGCATGTGGCACTGGGTGTGTGTGACGGGCTGC 1246
 Qy 1255 ACCCACTGGGCACTGTCGGGGGAGGAGCTG 1285
 Db 1247 ATCCGCTGGGCACTGTGTGACCGGCGCACTG 1277

 RESULT 12
 X25098
 ID X25098 standard; cDNA; 1277 BP.
 AC X25098:
 DE 05-JUL-1999 (first entry)
 DI Chicken Sonic hedgehog protein Shh cDNA.
 KW Sonic hedgehog; Shh gene; chicken; hedgehog therapeutic;
 KM ptc therapeutic; patched; chicken; signal transduction; muscle atrophy;
 KW cachexia; muscular myopathy; myoblastic sarcoma; therapy; ds.
 OS Gallus sp.
 PN W09J10004.A2.
 PF 04-MAR-1999.
 PR 28-AUG-1998; 017922.
 PR 29-AUG-1997; US-057394.
 PA (ONTO-) ONTOGENY INC.
 PI Bladen CS, Currie PD, Hughes SM, Ingham PW;
 DR WPT; 99-24357/20.
 P-PSDB: Y05510.
 PS A new method to regulate muscle growth
 PT Disclosure: Page 92-94; 130pp; English.
 CC This nucleotide sequence comprises a coding region for the chicken
 CC Sonic hedgehog protein Shh (see Y05510). The invention relates to
 CC a method for modulating the formation and/or maintenance of muscle
 CC tissue by ecotopically contacting muscle cells, especially muscle
 CC stem/progenitor cells, in vitro or in vivo, with a hedgehog
 CC therapeutic (i.e. hedgehog polypeptides and gene therapy
 CC constructs) or ptc therapeutic (i.e. a small organic molecule that
 CC mimics the effect of hedgehog proteins on patched signalling) or
 CC activates or potentiates patched signalling) in an amount effective
 CC to alter the growth state of the treated cells. Also claimed is a
 CC method for treatment or prevention of disorders of, or surgical or
 CC cosmetic repair of, such muscle tissues, by administering a
 CC hedgehog polypeptide or ptc therapeutic. The disorder may be
 CC muscle atrophy, in particular skeletal muscle atrophy or cardiac
 CC muscle atrophy, cachexia, or muscular myopathy (all claimed). The
 CC hedgehog polypeptide or ptc therapeutic can inhibit growth of
 CC myoblastic-derived tissue to provide treatment of hyperplastic or
 CC neoplastic growth of muscle tissue such as in myoblastic sarcoma
 CC (also claimed). The hedgehog therapeutic preferably comprises at
 CC least a bioactive extracellular portion of a hedgehog protein (see
 CC Y05510-19) encoded by a vertebrate hedgehog gene (see X25098-107),
 CC especially a human hedgehog gene.
 SQ Sequence 1277 BP; 262 A; 396 C; 387 G; 232 T;

Query Match

28.3%; Score 459; DB 1; Length 1277;

Best Local Similarity 65.2%; Pred. No. 3,9e-82;
 Matches 763; Conservative 0; Mismatches 360; Indels 48; Gaps 4;
 Qy 163 GAGCGCCCGCCAGCAAACTCGTGGCGCTGAGCTTCAAGCACTTACGCCCAATGTGCCG 222
 Db 107 GGAGGACCCCAAAAGCTGACCCCGTGTAGCTTAAGCACTTTATTCCTCAATGTGGAG 166
 Qy 223 AGAAGACCTTGGGCGCCAGCGGAGCGTATGAAGGCAAGATGCTGTCCAGCGCT 282
 Db 167 AGAAGACCTTGGGCGCGAGTGAAGATATGAAGGGAATATCAAGAAACTCCGAGAT 226
 Qy 283 TCAAGAGCTACACCCCAATTCATTCAGACATCATCTTCAAGAGACAGAGAAACAG 342
 Db 227 TTAAGAACTAACCCCAATTCACACCTTGACATTTATTTTAAGATGAAGAAACAGCG 286
 Qy 343 GCGCGGACCGGCTCATGACCGAGCGCTGCAAGACCGGCTGAGCTGTGTATCTCGG 402
 Db 287 GAGCTGACAGACTGATGACTACAGCGCTGCAAGACACACTGAATGCCCTGGCGATCGG 346
 Qy 403 TGATGAACCACTGGCCGCGGTGTGAAGCTGCGGGGTGACCGAGGGCTGGAGACAGCGCC 462
 Db 347 TGATGAACCACTGGCGCGGGGTGAAGCTGCGGGGTGACCGAGGGCTGGAGAGATGGCC 406
 Qy 463 ACGACTGAGAGAGATCCCTGCTATATGAGGCGCGGCGGTGGAATATCACATCAAGAC 522
 Db 407 ATCACTCGGAGGATCGGTGCTACAGAGGGTCCGCGCTGGACATCACAGTGTGGATC 466
 Qy 523 GCGACCGCAATAGATGATGACTGTGGCGCGCTTGGAGTGAAGCGCGCTTGTACTGG 582
 Db 467 GGGACCGGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 526
 Qy 583 TGTATTACGAGTCAAAAGCCCAAGTGTGCTCTGCTCAAGTCCGAGCACTCGCGCCAG 642
 Db 527 TCTATTAGAGATCCAAAGCGGACATCACTGCTCGCTCAAAAGCAAGAACTCAGTGGAG 586
 Qy 643 CCAAGACGGGGCGCTCTCTCTGCGGAGCCAGTACGCTGGAAGTGGGGCGGCTG 702
 Db 587 CGAAATCGAGAGGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 646
 Qy 703 TGGCTTGTACCGCTGTGAGCGCGGAGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 762
 Db 647 AGCTGTGAAGAGCACTGAGCGCTGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 706
 Qy 763 CCACCTTACGAGATGCTCATTTTCTGTAACCGGAGCCCAACAGCTGAGAGCTTCC 822
 Db 707 TGCTTACAGTACTCTCTCTCACTCTCTGACCGGATGAGACGCTCCGAAAGCTCTCT 766
 Qy 823 AGGTATGAGAGCTGAGAGACCGCCCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 882
 Db 767 AGGTATGAGAGCTGAGAGACCGCCCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 826
 Qy 883 TTACGGCTGACAAATACACAGGAGCGG-----CAGCCGCTTCCGAGCCACAT 930
 Db 827 TTGTGGCCCCCAGACACACAGTGGAGGACAGAGTCCACAGTGGGAGGCGCTCT 886
 Qy 931 TTGCGACCGACGTGACGCTGAGCA-----GTACGTCTGTGTGTGTGTGTGTGTGTGT 984
 Db 887 TCGCAGCAAGTGAAGGCTGCGCCCAAGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGT 946
 Qy 985 TCGACCTTCCCGCGGTGTGACAGCTGTCTACAC---ACGTGGCCCTGGGGGCTTACGCC 1041
 Db 947 TGTCTCCGCGGT 1006
 Qy 1042 CGCTCAAAAGATGGGACACTGT 1101
 Db 1007 CACTCACCCCGGAGGACCATCTCATACACGGGTGTGGCTCTCTCTCTCTCTCTCTCTA 1066
 Qy 1102 TGGTACACCACTGAGTGTGAGTGTGGCCCTTGAGACTCTTCAAGCTTGG 1161
 Db 1067 TCGAGAGACACAGTGTGGCCCATGTGGCCCTTGACACATCTCGCTTGGCTCAAGGGCTGC 1126
 Qy 1162 CATGGGCGACGTGAGCCCGGGGGAG-----GGTGTGC 1194

Dh 1127 TGGCCCGCTCTGCCCCAGATGGGGCCATCCCTACTGCCGCCACCAACCACTGGCATTC 1186
Qy 1195 ATTGTAACCCCACTGCTCTACCGCTGGGGCTCTCTGCTGAAGAAGGAGCACTTCC 1254
Dh 1187 ATTGTAACCAAGGCTCTCTACCGCATCGGACACTGGGTCTCTGATGGTACCGCGCTGC 1246
Qy 1255 ACCCACTGGGCACTGCTCGGGGAGGAGCTG 1285
Dh 1247 ATCCGCTGGGCACTGCTGGACCGGCGCAGCTG 1277

RESULT 13

X07270
ID X07270 standard; cDNA; 1190 bp.
AC X07270;
AT 21-MAY-1999 (first entry)
DE Human Desert hedgehog (Dhh) cDNA.
KW Desert hedgehog; Dhh gene; human; dopaminergic; GABA-nergic;
KW ptc therapeutic; patched; signal transduction; Parkinson's disease;
KW Huntington's disease; amyotrophic lateral sclerosis;
KW cerebral ischaemia; hypoxia; neuroprotective; gene therapy; ss; ds.
OS Homo sapiens.
PN M09904775-A2.
PD 04-FEB-1989.
PF 24-JUL-1998; U15419.
PR 24-JUL-1997; U5-900220.
PI (ONTO-) ONTOGENY INC.
PI Mahanthappa NK, Miao N, Pang K, Wang M;
DR WPI: 99-142578/12.
DR P-PSDB; W97764.
PT Increasing the survival of neuronal, dopaminergic and GABA-nergic
PT cells - by using a ptc therapeutic such as a protein kinase
PT inhibitor, or an agent derived from hedgehog polypeptides, useful in
PT the treatment of Parkinson's disease
PS Disclosure: Page 97-99: 138pp: English.
CC This nucleotide sequence comprises a coding region for the human
CC Dhh desert hedgehog protein (see W97764). The invention is based on
CC the finding that hedgehog proteins are useful as protective agents
CC in the treatment and prophylaxis of neurodegenerative disorders
CC resulting from the loss of dopaminergic and/or GABA-nergic neurons,
CC or the general loss of tissue from the substantia nigra.
CC Exemplary disorders include Parkinson's disease, Huntington's
CC disease (both claimed), amyotrophic lateral sclerosis and cerebral
CC ischaemia. The invention relates to hedgehog therapeutics (i.e.
CC hedgehog polypeptides and gene therapy constructs e.g. constructs
CC encoding recombinant hedgehog polypeptides and trans-activation
CC constructs for altering hedgehog gene regulatory sequences) and
CC ptc therapeutics (i.e. agents which mimic the effect of naturally
CC occurring hedgehog proteins on patched signalling) that are
CC effective in both human and animal subjects. A bioactive
CC polypeptide comprising amino acid residues 23-198 of human Dhh is
CC preferred. The products can also be used for the maintenance of
CC differentiated neurons in cultures, and to enhance the implantation
CC of such neuronal cells in an animal. They can also be used to
CC prevent or treat neurodegenerative conditions arising from the use
CC of certain drugs, and in the prevention and/or treatment of hypoxia,
CC e.g. as a neuroprotective agent.
SQ Sequence 1190 bp; 176 A; 375 C; 424 G; 215 T;

Query Match 28.0%; Score 453.8; DB 1; Length 1190;
Best Local Similarity 63.5%; Pred. No. 4.1e-81;
Matches 730; Conservative 0; Mismatches 407; Indels 12; Gaps 2;

Qy 103 TGTGCTGTGTGTGCTCCCGGCGATGGGGCTGGGGCGGGTGGTGGGACGCGCC 162
Dh 38 TGGCATCTTGTGGCTGGCCAGACCTGGGGCGGGCGGGCGGGTGGGCGCG 97
Qy 163 GGGACCGCACAGCA---AACTGCGCGCTGCGCTACAGCAAGTTAGGCCAATGTGC 219
Dh 98 GCGGCTATGGCGGACAGAGCTGCGCGCTACTACAAACAATTTTGGCGGCGTGC 157
Qy 220 CCGAAGAAAGACCTTGGGGCGGACGAGCTATGAAGCAAGATGCTGACGCTCGAGC 279

Dh 158 CAGAGCGGACCTTGGGGCGCAAGTGGGCGAGCGGAGGGAGGTGGCAAGGGGCTCCAGC 217
Qy 280 GCTTCAAGAGAGCTCAACCCCAATTAATCAGACATCATCTTCAAGAGAGAGACA 339
Dh 218 GCTTCCGGGAGCTCTGTCCTCCCACTCAACCCCGACATCATCTTCAAGAGAGACA 277
Qy 340 CAGGCGCGACCGGCTGATGACCGAGCGTGAAGGACCGGCTGAACTGCTGTATCT 399
Dh 278 GTGGAGCGCAACCGGCTGATGACCGAGCGTGAAGGAGAGGAGGAAAGCTTTGGCAATG 337
Qy 400 CGGTGATGAACCACTGGCCCGGTGTGAAGTGGGGGTGACCAAGGCTGGAGAGAGACG 459
Dh 338 CCGTATGAACATGTGGCCCGGAGTGGCTTACAGTGAAGTGAAGGCTGGAGAGAGC 397
Qy 460 GCGACCACTCAAGAGAGTCCCTGATATGAGGCGCGGGGGAGATCAACCAATCAG 519
Dh 398 GCGACCACTCAAGAGATTCCTCACTACGAGAGCGCTTGGACATCACTACGCTTG 457
Qy 520 ACCGCGACCGCAATTAATGATGAGTGGTGGCGGCTTGGAGAGAGCGGCTTGTACT 579
Dh 458 ACCGCGACCGCAATTAATGATGAGTGGTGGCGGCTTGGAGAGAGCGGCTTGTACT 517
Qy 580 GGGTGTATTACAGTCAAGAGCGGCGATGCTGCTGCAAGTCCGAGACTGGCGG 639
Dh 518 GGGTGTATTACAGTCAAGAGCGGCGATGCTGCTGCAAGTCCGAGACTGGCGG 577
Qy 640 CAGCGAAGAGAGGGGGGCTGCTCCCTCGGAGAGCGGCAAGTACGCTGAGAGTGGGGCG 699
Dh 578 GGGTGTATTACAGTCAAGAGCGGCGATGCTGCTGCAAGTCCGAGACTGGCGG 637
Qy 700 GTGTGCGCTTTCAGCCGCTGAGGCGGGAGACCGTGTGCTGGCCATGGGGAGAGTGGGA 759
Dh 638 GGAAGGGCTGGGGAGCTGACCGGAGACTGAGTGGTGGGGCGGAGAGCGTCAAGGCC 697
Qy 760 GCGCCACTTTCAGGATGTGCTCATTTCTCGACCGGAGCGGCGGAGCTGAGAGCT 819
Dh 698 GGGTGTATTACAGTCAAGAGCGGCGATGCTGCTGCAAGTCCGAGACTGGCGG 757
Qy 820 TCCAGTCAAGAGAGTCAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 879
Dh 758 TGTGTGCTGTGAGAGACCGAGTGGCTTCCAGCAAACTGTGCTACGCTTGGACTTG 817
Qy 880 TCTTACGGCTGACAAATCAACAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 939
Dh 818 TGTGTGCGGCTGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 877
Qy 940 ACGTGCAGCTGGCGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 999
Dh 878 GGTGTGCGGCTGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 937
Qy 1000 TGGGAGTGTCTTACACAGTGGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1059
Dh 938 TGGGAGTGTCTTACACAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 997
Qy 1060 CACTGTGTGAGAGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1119
Dh 998 GGTGTGCGGCTGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1057
Qy 1120 CTCAGTTGGCTTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1170
Dh 1058 GCGACCGCGGCTTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1117
Qy 1171 GCTGAGACCGGCGGAGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1230
Dh 1118 GGGCGGTCAAGCGCACTGGATGATGTGATCTCGGCTCTCTTACCGCTTACCGGAGG 1177
Qy 1231 TCTGTCTAG 1239
Dh 1178 AGCTACTGG 1186

-RESULT 14

X25105
ID X25105 standard; cDNA: 1190 BP.
AC X25105;
DT 05-JUL-1999 (first entry)
DE Human Desert hedgehog protein Dhh cDNA.
KW Desert hedgehog; Dhh gene; human; hedgehog therapeutic;
KW ptc therapeutic; patched; signal transduction; muscle atrophy;
KW cachexia; muscular myopathy; myoblastic sarcoma; therapy; ds.
OS Homo sapiens.
PN WO9910004-A2.
PD 04-MAR-1999.
PF 28-AUG-1998: U19922.
PR 29-AUG-1997: US-057394.
PA (ONTO-) ONTOGENY INC.
PI Bladgen CS, Currie PD, Hughes SM, Ingham PW;
DR P-PSDB; Y05517.
PT A new method to regulate muscle growth
PS Disclosure: Page 106-108: 130pp; English.
CC This nucleotide sequence comprises a coding region for the human
CC Desert hedgehog protein Dhh (see Y05517). The invention relates to
CC a method for modulating the formation and/or maintenance of muscle
CC tissue by ecologically contacting muscle cells, especially muscle
CC stem/progenitor cells, in vitro or in vivo, with a hedgehog
CC therapeutic (i.e. hedgehog polypeptides and gene therapy
CC constructs) or ptc therapeutic (i.e. a small organic molecule that
CC mimics the effect of hedgehog proteins on patched signaling, or
CC activates or potentiates patched signaling) in an amount effective
CC to alter the growth state of the treated cells. Also claimed is a
CC method for treatment or prevention of disorders of, or surgical or
CC cosmetic repair of, such muscle tissues, by administering a
CC hedgehog polypeptide or ptc therapeutic. The disorder may be
CC muscle atrophy, in particular skeletal muscle atrophy or cardiac
CC muscle atrophy, cachexia, or muscular myopathy (all claimed). The
CC hedgehog polypeptide or ptc therapeutic can inhibit growth of
CC myoblastic-derived tissue to provide treatment of hyperplastic or
CC (also claimed). The hedgehog therapeutic preferably comprises at
CC least a bioactive extracellular portion of a hedgehog protein (see
CC Y05510-19) encoded by a vertebrate hedgehog gene (see X25098-107),
CC especially a human hedgehog gene.
SQ Sequence 1190 BP; 176 A; 375 C; 424 G; 215 T;

Query Match 28.0%; Score 453.8; DB 1; Length 1190;
Best Local Similarity 63.5%; Pred. No. 4.1e-81;
Matches 730; Conservative 0; Mismatches 407; Indels 12; Gaps 2;

QY 103 TCCTGTGTGTCGCGCGGCGATGGGGCGGCGGTGGGCGTGGGACGCGCC 162
DB 38 TGGCACTTGTGGCGCTGCCAGCCAGAGCTGGGGCGGGCGGCGGTGGCGCGC 97
QY 163 GGCAGCCGCCAGCA---AACTGTGCGCTGCGCTACAGAGTTCAAGCCCAATGTC 219
DB 98 GCGCGATGCGGGCAAGCAGCTGCGCTACTCTACAGCAATTTTGGCCGGCGTGC 157
QY 220 CCGAGAGACCTTGGGCGCCAGCGAGCTATGAAGCAAGATTCGCTCGACCTCGAGC 279
DB 158 CAGAGCGGACCTTGGCGCGCCAGCGAGGGGAGGTGGCAAGGGCTCCGAGC 217
QY 280 GGTTCAGAGAGCTCAACCCCAATTACATCCAGACATCTCTTCAAGAGCGAGGAGACA 339
DB 218 GCTTCGGGAGCTCTGTCGCAACTACACCCGACATCTTCAAGAGTGAAGGAGACA 277
QY 340 CAGGCGCCAGCGCTCATAGCCAGCGGTGCAAGAGCCGTGAAGTGGTGGCTATCT 399
DB 278 GTGGAGCCAGCGCTCATAGCAGCGCTTGCAGAGAGAGGGTGAAGCTTTGGCCATTG 337
QY 400 CGGTATGAACAGTGGCGCGGTGTGAAGCTCGGGTGAACCGAGGGCTGGGAGAGAGC 459
DB 338 CCGTATGAACATGTGGCGCGGTGCGCTGAGAGTACTGAGGGCTGGGAGAGAGAGC 397
QY 460 GGCACCACTCAAGAGAGTCCCTGCAATTATGAGGGCGCGCGGTGGACATCACCAATCAG 519

DB 398 GGCACCACTCAAGAGAGTCCCTGCAATTATGAGGGCGCGCGGTGGACATCACCAATCAG 457
QY 520 ACCGCGCCAGCAATAGATGAGACTGTGGCGCGCTTGGAGTGAAGCGCGCTTGAAT 579
DB 458 ACCGCGCCAGCAATAGATGAGACTGTGGCGCGCTTGGAGTGAAGCGCGCTTGAAT 517
QY 580 GGGTATATTAGAGTCAAGAGCGCCAGGTGATGCTCCGTCAAGTCCAGAGACTCGGCCG 639
DB 518 GGGTCTACTACAGAGTCCGCAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 577
QY 640 CAGCAAGAGCGCGCGCTTCCCTGCCGAGAGCCAGGTGATCCCTGAGAGTGGGGCGC 699
DB 578 GGGTCCGGCGCGCGCTTCCCGGAAATGCAACTGTGCGCTGTGGAGAGGGCGAGC 637
QY 700 GTGTGCGCTTGTCAAGCGCTGAGAGCGCGGAGACCGGTGCTGCGCATGGGAGAGATGGA 759
DB 638 GGAAGAGGCTGGGGAGATGCAACCGCGGAGACTGGGTTTGGCGCGCATGCTCAGGCC 697
QY 760 GCCCACTTACAGAGTGTCTATTCTTCTGAGACCGCGAGCCCAAGCTGAGAGCTT 819
DB 698 GGGTGTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 757
QY 820 TCCAGTATGAGAGTCAAGAGACCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 879
DB 758 TTGTGTGTGTGAGAGCGAGTGGCTTCCAGCAAACTGTGCTGCTGCTGCTGCTGCTGCTG 817
QY 880 TCTTACAGCGTGAACATACAGAGCGCGGAGCGCGCTTCCGCGCATTTGCGACAGC 939
DB 818 TGTTCGCGCTGAGAGCG 877
QY 940 ACGTCAAGCTTGGCGAGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 999
DB 878 GGGTGTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 937
QY 1000 TGGCAGCTGTCTCTACAGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1059
DB 938 TGGCGCGTGTGGCG 997
QY 1060 CACTGTGTGTGAGAGATGT 1119
DB 998 GCGT 1057
QY 1120 CTGAGT 1170
DB 1058 CGCAGCGCGCTTGT 1117
QY 1171 GCTGAGCG 1230
DB 1118 GGGCGGTCCAGCG 1177
QY 1231 TCCTGTAG 1239
DB 1178 AGCTACTGG 1186

RESULT 15
O91639
ID O91639 standard; cDNA: 1425 BP.
AC O91639;
DT 08-MAR-1996 (first entry)
DE Human sonic hedgehog protein gene.
KW Human; sonic hedgehog gene; nested polymerase chain reaction; PCR;
KW fetal lung; probe; primer; diagnostic; nervous system disorder;
KW gene therapy; antibody; ss.
OS Homo sapiens.
PN WO951856-A1.
PD 13-JUL-1995.
PF 30-DEC-1994: U14992.
PR 30-DEC-1993: US-176427.
PR 14-DEC-1994: US-356060.
PA (HARD) HARVARD COLLEGE.
PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2000, 07:55:21 ; Search time 975.44 Seconds

(without alignments)
4948.935 Million cell updates/sec

Title: US-08-900-220-8

Perfect score: 1191
Sequence: 1 ATGGCTCCTCGACCAATCT.....CGAGAGACTACTGGGCTGA 1191

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_est48:*
83: gb_est49:*
84: gb_est50:*
85: gb_est51:*
86: em_est35:*
87: em_est36:*
88: em_est37:*
89: em_est38:*
90: gb_est52:*
91: gb_est53:*
92: gb_est54:*
93: gb_est55:*
94: gb_est56:*
95: em_est39:*
96: em_est40:*
97: em_est41:*
98: em_est42:*
99: em_est43:*
100: em_est44:*
101: em_est45:*
102: gb_est57:*
103: gb_est58:*
104: em_est46:*
105: gb_est59:*
106: gb_est60:*
107: gb_est61:*
108: gb_est62:*
109: gb_est63:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result

Query

No.	Score	Match length	DB	ID	Description
1	254	21.3	314	28	AA064660 zml3f08.s
2	239	20.1	357	49	AI645932 mml1g06.y.y
3	234.8	19.7	463	49	AI646532 mml1g06.x
4	171.6	14.5	299	69	AM144802 EST291875
5	121.8	10.2	406	63	AI958076 f69e0l.y
6	72.6	6.1	932	82	AI066742 Drosophila
7	71	6.0	459	41	AI015155 Drosophila
8	70.4	5.9	459	41	AI015155 Drosophila
9	70.2	5.9	925	82	CNS0091P
10	67.8	5.7	512	30	AA245525 my52c03.r
11	67	5.6	352	27	AI325370 m134b11.y
12	65.4	5.5	434	24	AA032692 m134b11.r
13	63.8	5.4	935	82	CNS006XK
14	63	5.3	1101	83	CNS0175X
15	62	5.2	932	82	CNS0072Q
16	61.6	5.2	935	82	CNS0091P
17	61.6	5.2	925	82	CNS006XK
18	59.8	5.0	776	82	CNS010RY
19	58.8	4.9	839	82	CNS004NB
20	56.6	4.8	1101	82	CNS001XT
21	56	4.7	1203	83	CNS015X4
22	55.4	4.7	970	82	CNS010C9
23	55.2	4.6	1203	83	CNS015Y4
24	54.4	4.6	1201	83	CNS016EM
25	53.6	4.5	915	83	CNS015YX
26	53	4.5	991	74	AF122121
27	52.2	4.4	1101	83	CNS004XP
28	51.8	4.3	977	82	CNS001K7
29	51.8	4.3	1009	82	CNS010EW
30	51.8	4.3	1009	82	CNS010EW
31	51.6	4.3	1101	83	CNS0153F
32	51.2	4.3	910	82	CNS006ON
33	50.8	4.3	955	82	CNS006SE
34	50.4	4.2	1101	82	CNS00397
35	49.8	4.2	645	82	CNS012I3
36	49.2	4.1	1036	82	CNS010BS
37	49	4.1	1077	83	CNS0156Q
38	48.8	4.1	1101	83	CNS0181I
39	48.6	4.1	692	82	CNS007WH
40	48.6	4.1	844	82	CNS0052P
41	48.6	4.1	962	82	CNS00K3S
42	48.6	4.1	1101	83	CNS016HG
43	48.4	4.1	1171	84	B10823
44	48.2	4.0	918	91	AB089329
45	48	4.0	834	82	CNS00KXV
					AA064660 zml3f08.s
					AI645932 mml1g06.y.y
					AI646532 mml1g06.x
					AM144802 EST291875
					AI958076 f69e0l.y
					AI066742 Drosophila
					AI015155 Drosophila
					AI015155 Drosophila
					AI010855 Drosophila
					AI035013 Drosophila
					AA245525 my52c03.r
					AI325370 m134b11.y
					AA032692 m134b11.r
					AI066450 Drosophila
					AI066742 Drosophila
					AI055013 Drosophila
					AI065051 Drosophila
					AI099332 Drosophila
					AI054280 Drosophila
					AI076875 Drosophila
					AI068054 Drosophila
					AI096787 Drosophila
					AI106054 Drosophila
					AI106648 Drosophila
					AI106083 Drosophila
					AF122121
					AI106440 Drosophila
					AI076850 Drosophila
					AI076882 Drosophila
					AI096882 Drosophila
					AI096882 Drosophila
					AI104949 Drosophila
					AI065629 Drosophila
					AI065794 Drosophila
					AI065921 Drosophila
					AI015589 Drosophila
					AI096770 Drosophila
					AI105068 Drosophila
					AI105068 Drosophila
					AI0505923 Drosophila
					AI0505652 Drosophila
					AI077188 Drosophila
					AI106750 Drosophila
					B10823
					AI078329
					HS_4832_A
					AI078276 Drosophila

ALIGNMENTS

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
AA064660	LOCUS	AA064660	314 bp	MRNA	EST	23-DEC-1997		
		zml3708.s1	Stratagene pancreas (#937708)	Homio sapiens cDNA clone				
		IMAGE:555543	3 similar to TR:0443942	G443942	DESERT	HEDEHOG		
		PRECURS0R	;	mrna sequence.				
		AA064660						
		AA064660.1	GI:1558743					
		EST.						
		human.						
		Homio sapiens						
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;						
		Eutheria; Primates; Catarrhini; Homiidae; Homo.						
		1 (bases 1 to 314)						
		Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chappell,B.,						
		Chissoe,S., Dietrich,N., Dubouque,T., Favelllo,A., Gish,W.,						
		Hawkins,M., Hultman,M., Kueba,T., Lacy,M., Le,M., Le,N.,						
		Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,						
		Ronfing,T., Schellenberg,K., Soares,W.B., Tan,F., Thierry-Mieg,J.,						
		Trevaskis,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R.						
		and Marra,M.						

TITLE	Generation and analysis of 280,000 human expressed sequence tags.
JOURNAL	Genome Res. 6 (9), 807-828 (1996)
MEDLINE	97044478
COMMENT	On May 9, 1995 this sequence version replaced g1:802645. Contact: Wilson RK Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel.: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu WARNING: There is evidence that suggests that the 384-well parent plate of this clone contains both human and mouse derived clones. Thus, the origin of this clone is uncertain. This caution should be kept in mind should you use this clone.
FEATURES	This clone is available royalty-free through LINT; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Possible reversed clone: similarity on wrong strand Possible reversed clone: polyT not found Insert Length: 862 Std Error: 0.00 Seq primer: -40M13 fwd. from Amersham High quality sequence stop: 283. Location/Qualifiers
SOURCE	1. 314 /organism="Homo sapiens" /db_xref="GDB:3916972" /db_xref="taxon:9606" /clone="IMAGE:525543" /clone_id="Stratagene pancreas (#937208)" /lab_host="SOBR cells (kanamycin resistant)" /note="Organ: pancreas; Vector: pBluescript SK-. Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: 0190 dt. Pancreatic adenocarcinoma cell line. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGCACAGC 3' -3' adaptor sequence: 5' CTCAGTTTCTTTTTTTTTTTT 3'
BASE COUNT	31 a 100 c 117 g 55 t 11 others
ORIGIN	
Query Match	21.3%; Score 254; DB 28; Length 314;
Best Local Similarity	91.3%; Pred. No. 2.8e-40;
Matches 285; Conservative	0; Mismatches 25; Indels 2; Gaps 2;
OY	833 GGCCGGCCCCCGCCGACAGCACTTGACACCGGTTCGCCGCCGCCGCTACGCCGTGGG 892 Db 1 GGCCGGCCMCOCGGCAGNGNANNTTGCACCGGTTCGCCGCCGCCGCTACGCCCTNNGGG 60
OY	893 ACTCGGTCTGGCGCCCGCGGGGGGATGGCGCTTCGGCCACAGCGCGCGTGGCCGTGGCGC 952
Db	61 ACTCGGTCTGGCGCCCGCGGGGGATGGCGCTTCGGCCACAGCGCGCGTGGCCGTGGCGC 120
OY	953 GGGAGAAGCGGTGGCGGTTCGGCGCGGCTCACCGCGACGGAGCGGTGCTGGGAAGC 1012 Db 121 GGGAGAAGCGGTGGNGTTCGGCGCGCTCACCGCGACGGAGCGGTGCTGGGAAGC 180
OY	1013 ATGTCTGTGACCTCTTGCTACACGGGTTCGAGAGTACACAGTAGGAGCGACCGCGCTTTTG 1072
Db	181 ATGTCTGTGACCTCTTGCTACACGGGTTCGAGAGTACACAGTAGGAGCGACCGCGCTTTTG 240
OY	1073 CCCCCTTAGACTGTGTGACAGCGGTAGGGGGGTCTCCCGCGGGGGCGCTCCAGCGGA 1132 Db 241 CCCCCTTAGACTGTGTGACAGCGGTAGGGGC-NTCCTCCCGGC-GGGCCGTCCAGCGGA 298
OY	1133 CTGGCATGCATT 1144 Db 299 CTGGATGCATT 310
RESULT	2
A1645932	537 bp mRNA EST 29-APR-1999
DEFINITION	mullig06.y1 Soares_thymus_2nbmt Mus musculus cDNA clone IMAGE:639130


```

ACCESSION      A1645932
VERSION        A1645932.1
KEYWORDS       GI:4724407
SOURCE         EST.
ORGANISM       house mouse.
               Mus musculus.
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
               Eutheria; Rodentia; Sclurognathii; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 537)
AUTHORS        NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
               Tumor Gene Index
JOURNAL        Unpublished (1997)
COMMENT        On May 18, 1998 this sequence version replaced gi:3137802.
               Contact: Robert Strausberg, Ph.D.
               Tel.: (301) 496-1550
               Email: Robert.Strausberg@nih.gov
               This clone is available royalty-free through LNL, contact the
               IMAGE Consortium (info@image.llnl.gov) for further information.
               This read is a RESEQUENCE of a previously sequenced mouse clone
               This read has been verified (found to hit its original self in the
               correct orientation)
               Possible reversed clone: similarity on wrong strand
               Seq primer: -40RP from Gibco
               High quality sequence stop: 478.
               Location/Qualifiers
               1..537

```

Query Match	Best Local Similarity	Match	Conservative	Score	DB	Length	Mismatches	Indels	Gaps
1	ATGGGCTCTCTGACCAATCTACTGCGCTTGCTGCTGGATGGACACTTGTGCGGTGCGAGCC	60							
235	ATGGCTCTGCGCGCCAGTCTGTGCGCCCTGTCTGCTTGGCATCTGTGACATATCTGCG	294							
61	CAGAGCTGCGGGCCGGGCGGGGCGGCTTGCGCGGCGCGCTATGCGCGAAGCAGCT	120							
295	CAGAGCTGCGGGCCGGGCGCGAGAGCGAGTGGCGCGGCGCGCTTATGTGCGAAGCAACTT	354							
121	GTGGCGGCTACTCTATAAGCAATTTGTGCGCGCGCGCTGCGAGACCGACCTTGGGGCGCACT	180							
355	GTGGCTCTGCTATTAACAGCAGTTTGTGCGCGCAATATGCGCGAGCGACCTGTGGCGGAGT	414							
181	GGGCGCAGCGGAGGAGGCTGCGCAAGGGGCGCTCGAAGCGCTTCGCGGACCTCTGTGCCAAC	240							
415	GGGCGCAGCGGAGGAGGCTTACAAAGGGGCGTGGAGCGCTTCGCGGACCTCGTACCCAAC	474							
241	TACAAACCCGACATCACTTCAAGATGAGAGAACATGAGACCGACCGCGCTGATGACC	300							

Db 475 TACAAACCCGACATAATCTTCAAGAGTGGAGAGAAACAGCGCGCAGACCGCTGATGACA 534
 OY 301 GAG 303
 Db 535 GAG 537
 RESULT 3
 A1666359/c
 LOCUS
 DEFINITION
 3' similar to TR:Q61488 O61488 DESERT HEDGEHOG HOMOLOG PRECURSOR ;,
 mRNA sequence.
 A1666359
 A1666359.1 GI:4804713
 EST.
 house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 463)
 NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 On May 18, 1998, this sequence version replaced gi:3136856.

Other_Estis: muliguo.yl
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov

This clone is available royalty-free through LNLN; contact the IMAGE Consortium (infoimage.lnl.gov) for further information. This clone was previously sequenced on the 5' end only, this new data is from the 3' end

Possible reversed clone: similarity on wrong strand
High quality sequence stop: 261.

Location/Qualifiers

1..463

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone-"IMAGE:639130"
/clone_id="Soares_thymus_2NBWT"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"

/note="Vector: pT7AD-Pac (Pharmacia) with a modified polylinker. Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer 15', TGTTCACATCTGAAGCGGAGGGCGGCCTTTTCTTTTTTTTTTTTTTTT 3'); double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7A3 vector. RNA provided by Dr. Bertrand Jordan. Library went through two rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 87 a 124 c 138 g 113 t 1 others

ORIGIN

Query Match 19.7%: Score 234.8; DB 49; Length 463;
Best Local Similarity 85.8%: Pred. No. 1.6e-36;
Matches 260; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1 ATGGCTCTCTGACCATACTACTGCGCCCTTGTCGTCCTTGACACTTCTGGCGCTGCCAGCC 60
 | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 463 ATGGCTCTGCGGGCCAGAGCTGTGGCCCTGTGCTCTTGACACTTCTTGCACTATCTGCC 404
 | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 61 CAGAGCTCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGCAAGAGAGCTC 120
 | | | | | | | | | | | | | | | | | | | | | | | | | | |

DB 403 CAGAGCTCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGCAAGCAACTT 344

OY		121	GCGGGCTACTTACAAAGATTTTGCCGCCCCGCGRGCCAAGACCAGCAGCATTCGTGGCCCACT	180
Dd		343	TGCTCTTGCTATTACAAGCATTTTGTCCTATAATGCCCAGCGGAACCTTGCCCGTAGT	284
OY		181	GCGCCAGCGGAGGAGGGTGGCAAAGGGGCTCCGAGCGCTTCCGGACCTCTGCTCCAAC	240
Dd		283	GCGCCAGCGGAGGAGGGTGGTAACAGGGGGGTGGAGCGGTTTTCCGGACCTCTGTAACAAC	224
OY		241	TACAACCCCGACATATCTTCAAGATGAGAAGAACAGTGCAGCCGACCGCTATGACC	300
Dd		223	TACAACCCCGACATATCTTCAAGATGAGAAGAACAGCGCGCACCGCTGTATGACA	164
OY		301	GAG 303 	
Dd		163	CAG 161	
RESULT		4		
LOCUS		AU144802		
DEFINITION		AWI44802	299 bp mRNA EST	30-OCT-1999
ACCESSION		ESR291875	Normalized rat embryo, Bento Soares Rattus sp. cDNA clone	
VERSION		RGIIGS5	5' end similar to Indian hedgehog protein, mRNA sequence.	
KEYWORDS		AWI44802.1	GI:6161619	
SOURCE		EST.		
ORGANISM		Rattus sp.		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
AUTHORS		Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. Lee,N.H., Glodex,A., Chandra,I., Mason,T.M., Quackenbush,J., Kerlavage,A.R. and Adams,M.D. Rat Genome Project: Generation of a Rat EST (RESE) Catalog & Rat Gene Index		
JOURNAL		Unpublished (1998)		
COMMENT		On Jun 5, 1998 this sequence version replaced gi:3189461.		
		Contact: Lee, NH		
		AtNC		
		The Institute for Genomic Research 9712, Medical Center Drive, Rockville, MD 20850, USA Tel.: (301)-838-3529 Fax: (301)-838-0208 Email: nhlee@tigr.org		
		For clone availability, additional sequence and expression information related to this EST please check the TIGR Rat Gene index (http://www.tigr.org/tcld/rgl/rgi.html). To order a clone contact the ATCC (http://www.atcc.org/atcc.html). Seq primer: M13 Reverse.		
FEATURES		Location/Qualifiers		
source		1..299	/organism="Rattus sp."	
		/db_xref="taxon:10118"		
		/clone="RGICG52"		
		/clone_1lb="Normalized rat embryo, Bento Soares"		
		/dev_stage="embryo 8, 12, 18 dpc"		
		/note="vector: pMT3pac; Site_1: EcorI; Site_2: NotI"		
BASE COUNT		78 a 92 c 80 g 49 t		
ORIGIN				
Query Match		14.5% Score 172.6 DB 69 Length 299; Best Local Similarity 73.6%; Pred.No.1.7e+24; Matches 220 Conservative 0 Mismatches 79 Indels 0 Gaps 0;		
OY		177	CAGTGGCCAGCGGAGGGAGGAGGATGCAAGGGCTCCGAGCGCTTCCGGACCTGTGACC	236
Dd		1	CAGCGGGGCTACTTAAGCAAAAATGCGCGCAGCTCTAGCGCTTCAAGAGACTACCCC	60
OY		237	CAACTACAACCCCGACATCATCTTCAAGANTGAGAGAACAAGTGGAGCGGACCGCTGAT	296
Dd		61	CAACTACAATCCCGACATCATCTTCAAGAGACAGAGAACAACCGGTGCCGACCGCTCAT	120
OY		297	GACCAAGGCTTCAAGGAGAGGGGTAAAGCGTTTGGCCATTCGCTGATGATCATGTGCC	356

Db 121 GACCAGGCGCTGCAGAGGACCGCTTCACCTACGTGCGCATCTCTGTATCAATGAACCAATGGCC 180
OY 357 CGGAGTGGCGCTACAGTACTGAGGGCTGGAGACAGAGCGGACCAACCGTCAAGATTTC 416
Db 181 GGGTGTGAAGCTGGCGGGTACGACCGAGGCTGGGATGTAAGACCGGCATCACTCAGAGGAAATC 240
OY 417 ACTGCATCAACGAGGCCCTGCTTTGGATATCACTACGCTCTGACCGGACCGCACCAACT 475
Db 241 TTTCACATGATGAGGCGCGGGGTGGATATCAACCACTGACAGCGGACGAAATTAAGT 299

RESULT 5
LOCUS AI958076 406 bp mRNA EST 20-AUG-1999
DEFINITION fc90e01.y1 zebrafish wshu mrimg EST Danio rerio cdna 5' similar to
AI958076
ACCESSION AI958076
VERSION AI958076 GI:5750785
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
Cyprinoidae; Cyprinidae; Rasbora; Danio.
1 (bases 1 to 406)
Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Maria,M.,
Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theisling,B., Allen,M., Bowers,Y.,
Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., Mccann,R.,
Waterson,R. and Wilson,R.,
Wshu Zebrafish EST Project 1998
Unpublished (1998)
On Jun 5, 1998 this sequence version replaced gi:3188836.
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zdrafish@wustl.edu
CDNA Library Preparation: Matthew Clark, CDNA Library Arrayed by:
Matthew Clark, DNA Sequencing by: Washington University Genome
Sequencing Center Clone distribution: Genome Systems, St. Louis,
Missouri (web address: www.genomesystems.com) (email contact:
info@genomesystems.com) and Research Genetics, Huntsville, Alabama
(web address: www.resgen.com) (email contact: info@resgen.com) and
Resourcenzentrumumrumpfarmdatenbank, Berlin, Germany (web address:
www.xrpd.de)
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: T3 ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. 406
/organism="Danio rerio"
/db_xref="taxon:7935"
/clone_id="Zebrafish Wshu MPRMG EST"
/sex="mixed"
/tissue-type="26 somite embryos, adult livers, shield
stage embryos"
/lab_host="Xtl-blue MRF"
/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; 1st
strand cDNA was primed with a Not I - oligo(dT)15 primer
[5'pGATGATGTCAGATGCGGAGCGGCGCCCTTTTATTTTTTTTTTTT];
double-stranded cDNA was ligated to Sal I adaptors (pRL),
digested with Not I and cloned into the Not I and Sal I
sites of the pSPORT1 vector (pRL). Library was constructed
by Matthew Clark (Lehrach lab, ICRF, London and Max Planck
Institut fuer Molekulare Genetik, Berlin). cDNAs for EST
analysis were selected following oligonucleotide
hybridization fingerprinting of arrayed clones from
zebrafish late somitogenesis (26 ss), adult liver or

FEATURES	SOURCE	Location/Qualifiers
		1. 459
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/clone="IMAGE:1622169"
		/clone_lib="Soares_total_fetus_Nb2HF8_9w"
		/dev_stage="8-9 weeks"
		/lab_host="DH10B"
		/note="vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I - oligo(dT) primer [5' TGGTACCACTGAGGAGGAGCGGCGGCGCTAATTTTTTTTTTT 3'].
		Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT	106 a	127 c 133 g 93 t
ORIGIN		
Query Match	6.0%;	Score 71; DB 41; Length 459;
Best Local Similarity	67.1%;	Pred. NO. 0.0001;
Matches 116;	Conservative	0; Mismatches 55; Indels 2; Gaps 14
OY	82	GGGCGGTTGGCCGGCGCGCTATGGCGCAA--GCAGCTGTGCCGTACTCTACAGC 139
Db	454	GGGTGGGGTGGCGCGCGCGCGCGCAACCTCGTGGCGGCTCGGCTACAGC 395
OY	140	AATTGTGGCGGCGTGGCGAGAGCGGAGCCCTGGGGCGCATGGGCGAGCGAGGGAGG 199
Db	394	AGTTAGCCCCCATGTGCCCCGAGAAAGCCCTGGGGCGCGAGCGAGCTATGAGGAGAA 335
OY	200	TGGCAAGGGGCTTCGAGCGCTTCGCGGAGCCTGTGGCCCACTACATCAACCCGAG 252
Db	334	TGCGTCGGAGCGTCCGAGCGCTTCATGAGCTACACCCCAATTACATGCACAC 282
RESULT	8	
CNS016KD		
LOCUS		
DEFINITION	CNS016KD 1100 bp DNA GSS 26-JUL-1999	
	Drosophila melanogaster genome survey sequence SP6 end of BAC	
	BACN16D22 of DrosBAC library from Drosophila melanogaster (fruit	
	fly), genomic survey sequence.	
ACCESSION	AL106855	
VERSION	AL106855.1 GI:5624152	
KEYWORDS	GSS.	
SOURCE		
ORGANISM	fruit fly	
	Drosophila melanogaster	
	Euarysta, Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	
	Pterygota, Neoptera; Endopterygota; Diptera; Brachycera;	
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
REFERENCE	1 (bases 1 to 1100)	
AUTHORS	Genoscope.	
TITLE	Direct Submission	
JOURNAL	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr	
	- Web : www.genoscope.cns.fr)	
COMMENT	dermatation of this BAC-end and sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP)-http://www.edgp.edi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBAC11.	
FEATURES		
SOURCE	Location/Qualifiers	
	1. 1100	
	/organism="Drosophila melanogaster"	
	/plasmid="pBelOBAC11"	
	/db_xref="taxon:7227"	
	/clone_lib="DrosBAC"	

```

/closure="BACN16D22"/
/note="end : SP6"
BASE COUNT      132 a       229 c       106 g       220 t       413 others
ORIGIN

Query Match          5.9%   Score 70.4; DB 83; Length 1100;
Best Local Similarity 24.4%; Pred. No. 0.00016;
Matches 126; Conservative 175; Mismatches 213; Indels 2; Gaps 1;

Oy  616 GTGCCCTGTGAGCGAGCGAGCAAAAGGGCCGCAGCAACTCACCAGCGAGACATCGTT 675
    ||:::||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  566 GTSSGGSKGTGGTGTCCTCCGTSGGSGCGGTGTCTTAKSKGTGCGCBGGSSGCKSTGC 645
Oy  676 TTGGCGCCCATGATGCTCAGAGCCCGGGTGTGTGCCACAGCGCGGTGCTCTCTCTCGAACCGG 735
    ||:::||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  646 GSGSGGGGYGKSKSGSGCGSGSGSGSCGCSGSSSGCGKCGCSRKSGVSCSBBSGG 705
Oy  736 GACTTGACAGCGCGGGGCTTCAATTGTGGCTGTGGAGACGAGTAGTGCCCTCCACGAAATG 795
    ||:::||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  706 YCCBBGSSSKCCSBBGGGSGCGCGGCCGCCGTCGCCKBGBBGGSYGYCGGSKCB 765
Oy  796 TTGCTCACAGCCCCCTGGACACTGTGTGTTGCGCCTCAGAGGGCGGCGCCGCGCAGAGCAC 855
    ||:::||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  766 CTGKTCCCSCTSYTKKKSKCGGTGTGTCTCKKCCTGTCTCKBYCKCYTT--CKSYCK 823
Oy  856 TTTGCACCGCGATGTTCCGCGCGCGGCTACGCGCTGGGAGCTGGTGGTCGCGCCGCGG 915
    ||:::||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  824 BCYKCGCGCGCGSCBKSGCCSKCTCKTKCKYKKKKKYKBYKKYKCKCKTYKCKKBC 883
Oy  916 GATGGGCTTCGGCCAGCGCGCGTGGCCGTGTGGGCGGCGGAGAAAGCCGTGGCGTTC 975
    ||:::||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  884 KYKCKBKBCCKCKBKCKCKCKCKCKBKCBKCBKCKCKCKCBKCBKCBCKCYK 943
Oy  976 GCGCGCTCACCGCGCAGGGAGCGGTGCTGGGAAGATGTCGCTCTGTCAACGG 1035
    ||:::||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  944 CCKBCBCKCKYKCKCKBKBYKCBEBCKCCKCCBCKBCKCKBKCKCKCKCK 1003
Oy  1036 GTTCTGAGAAGTCACCAATGGGCGCAGCCGCTTTTGCCCTTGAAGTGTGTCAGTCGACGCG 1095
    ||:::||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  1004 CBKCBKBCBCKBKCKCKBCKBCKCKKYYTBKYYKKKKKKBTBKKKBYKCKYC 1063
Oy  1096 CTAGGCGCGCTGCTCCC CGCGGGGCGCGTCCAAGCC 1131
    ||:::||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  1064 BSCBKCKBKBCCKCBCKCBKCCCBCKBCK 1099

RESULT  9
LOCUS      CENS091P           925 bp     DNA              GSS               03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TENG end of BAC #
            BACR1D16 of RPc1-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION  AL053013
VERSION    AL053013.1 GI:4934461
KEYWORDS   GSS.
SOURCE     Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1 (bases 1 to 925)
AUTHORS   Genoscope.
TITLE     Direct Submissions
JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ;
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
COMMENT   - Web : www.genoscope.cns.fr)
            Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutayo Osoe-gawa and

```


REFERENCE	Eutheria: Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Euarchonta; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ATTHORS	1 (bases 1 to 452)
	Matta,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucada,T., Lacy,M., Le,M., Martin,U., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE	The WashU-HM1 Mouse EST Project
JOURNAL	Unpublished (1996)
COMMENT	On Jan 19, 1998 this sequence version replaced gi:2152292. Contact: Matta/Mouse EST Project WashU-HM1 Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LML ; contact the IMAGE Consortium (infoimage.lml.gov) for further information. MGI:279221
FEATURES	This read is a RESEQUENCE of a previously sequenced mouse clone This read has been verified (found to hit its original self in the correct orientation)
SOURCE	Seq primer: -40RP from Glbco High quality sequence stop: 415. Location/Qualifiers
	1..452
	/organism="Mus musculus"
	/strain="C57BL/6J"
	/db_xref="taxon:10090"
	/clone="IMAGE:465405"
	/clone_lib="Soares mouse embryo NMEL3.5 14.5"
	/sex="unknown"
	/issue_type="embryo"
	/dev_stage="13.5-14.5dpc total fetus"
	/lab_host="DH10B"
	/note="Vector: p773D-Pac (Pharmacia) with a modified polylinker. Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer 15' TGTTACCAATCTGAAGTGGGAGCGGCCCGGAAATTTTTTTTTTTTTTTTTTTT T 3'1, on equal amounts of mRNA from 2 13.5dpc and 2 14.5dpc embryos (total RNA provided by Minoru Ko, Wayne State Univ., from 2) ; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT773 vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT	70 a 137 c 140 g 105 t
ORIGIN	
Query Match	5.6%; Score 67; DB 44; Length 452;
Best Local Similarity	59.0%; Pred. No. 0.00061;
Matches 115, Conservative	0; Mismatches 80; Indels 0; Gaps 0;
07	888 GGGCGGGGGGGGGGAGATCGCTTGGCCAGCGCGCGCCCGCTTGGCGGGGAG 957
Db	1 GGTGGTATATCGGGGTATACAGAGCTCCAGCTGTCTGGGTGGAGATCTTCACCC 60
07	958 GAAGCGGTGGAGCGTTCGCGCGCGCTACACCGGACGAGCACTCTGTGAACATGC 1017
Db	61 GGGCGCTTGGTCTCTATCTCTCTCACAAAGCATGGACACTTGTGGAGATG 120
07	1018 CTGGCGCTTGGCTACGCGGTCTGGAAGATCACCAAGTGGCGCACCGGCTTTGGCCCC 1077
Db	121 GGGCGCTCTCTGTTCAGACTGTGGCTACACCAACTGTGCTCAGTTGGCTTGGCCC 180
07	1078 TTGAGATCGCTGCAC 1092
Db	181 CTGGCAGTGTCTTCCC 195

RESULT	12
AA032692	EST 22-AUG-1996
LOCUS	AA032692 324 bp mRNA
DEFINITION	m34b1.1 Soares mouse embryo NBME13.5 14.5 Mus musculus cDNA
ACCESSION	AA032692
KEYWORDS	protein - mouse ; mRNA sequence.
ORGANISM	AA032692.1 GI:1505275
SOURCE	EST.
REFERENCE	house mouse.
AUTHORS	Mus musculus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
	1 (bases 1 to 324)
	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
	Geisel,S., Kucada,T., Lacy,M., Le,M., Martin,J., Morris,M.,
	Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
	Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
	Waterson,R.
TITLE	The Mashu-HMI Mouse EST Project
JOURLAL	Unpublished (1986)
COMMENT	On Nov 29, 1993 this sequence version replaced g1:637531.
	Contact: Marra M/Mouse EST Project
	Mashu-HMI Mouse EST Project
	Washington University School of Medicine
	444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
	Tel: 314 286 1800
	Fax: 314 286 1810
	Email: mouseest@watson.wustl.edu
	This clone is available royalty-free through LINT ; contact the
	IMAGE Consortium (info@image.llnl.gov) for further information.
	GI:279221
	Possible reversed clone: similarity on wrong strand
	Seq primer: -28M13 rev2 from Amersham
	High quality sequence stop: 98.
FEATURES	Location/Qualifiers
source	1..324
	/organism="Mus musculus"
	/strain="C57BL/6J"
	/db_xref="taxon:10090"
	/clone="IMAGE:465405"
	/clone_id="Soares mouse embryo NBME13.5 14.5"
	/sex="unknown"
	/issue_type="embryo"
	/dex_stage="13.5-14.5dp total fetus"
	/lab_host="DH10B"
	/note="Vector: pT73D-Pac (Pharmacia) with a modified
	polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
	was primed with a Not I - oligo(dt) primer [5'
	TGTACCAATCTGTAAGTGGGCGCGCGGGAATTTTTTTTTTTTTTTTTT
	T 3'], on equal amounts of mRNA from 2.13 dp and 2
	14.5dp embryos [total RNA provided by Minoru KO, Wayne
	State Univ., from 2 J]; double-stranded cDNA was ligated to
	Eco RI adaptors (Pharmacia), digested with Not I and
	cloned into the Not I and Eco RI sites of the modified
	pT73 vector. Library went through one round of
	normalization, and was constructed by Bento Soares and
	M.Felima Bonaldo."
BASE COUNT	49 a 96 c 100 g 79 t
ORIGIN	
Query Match	5.5%; Score 65.4; DB 27; Length 324;
Best Local Similarity	58.5%; Pred. No.0.0011;
Matches 114; Conservative	0; Mismatches 81; Indels 0; Gaps 0;
898 GTGCTGGCGCCCGGCGGGAGATGCTTGCGCCACGCGCGTGGCCCTGTGGCGGGAG	957
Db 1 GTGCTGATATGAGGGTATCAAGCCTTCCAGCTGTGTGGGTGACAGCTGTCTCCACCC	60
958 GAAAGCGTGGCGGTTCGCGCCCTCAACCGCGACGGAGCGTCTGTGTGAACGATGTC	1017

```

Db 61 GTGGCCCTTGAGTCTATGCTCTCTCAGACAGCATGGACACTTGATGAGATG 120
QY 1018 CTGGCCCTTGCTACCGGATTCGTGAGAGATCACAGTGGCGCACCGCTTTGCCCC 1077
Db 121 GTGGCCCTTGCTGAGCTGTGGCTGACACACATCTGCTCACTTGCTTGGCCAC 180
QY 1078 TTGACATCTGCTAC 1092
Db 181 CTGCACCTGTTCCC 195

RESULT 13
CNS006X/c 935 bp DNA GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC #
DEFINITION BACN14N09 of RPCL-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL066051 GI:4945019
VERSION AL066051.1 GI:4945019
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 935)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNML Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Oseegawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCL-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
Source location/Qualifiers
1..935
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCL-98"
/clone="BACN14N09"
/note="end : T7"

BASE COUNT 257 a 170 c 162 g 96 t 250 others
ORIGIN
Query Match 5.4%: Score 63.8; DB 82; Length 935;
Best Local Similarity 32.3%; Pred. No. 0.003;
Matches 123; Conservative 90; Mismatches 164; Indels 4; Gaps 1;

```

```

QY 980 CGCTACCCGCGACGAGCGCTGCTGTGAACGATGCTCTGACCTTACGCGGTTTC 1039
Db 755 GSSCGSSSSGSGSGCGGSCSSCGSSGSGSGCGCGCCGCGCGCGCGSS 696
QY 1040 TGGAGATCACCAGTGGGCGCACCGCGCTTTGGCCCCCTTAGANTGCTGACCGGCTAG 1099
Db 695 GCGSGCGSGSGCGCGCGCGCCCGCCCGCCSSCGSSCGSSCGSSCGSSCGCC 636
QY 1100 GAGCGCTGCTCCCGCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1159
Db 635 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 576
QY 1160 TCTACCGCTTAGCGGAGAGC 1180
Db 575 GSASGCGCGCGGCGRAGSGKM 555

RESULT 14
CNS017SY/c 1101 bp DNA GSS 26-JUL-1999
LOCUS Drosophila melanogaster genome survey sequence SP6 end of BAC
DEFINITION BACN37L08 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL108460 GI:5628764
VERSION AL108460.1 GI:5628764
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNML Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billand at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

FEATURES
Source location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/plasmid="pBelobAC11"
/db_xref="taxon:7227"
/clone_lib="DrosBAC"
/clone="BACN37L08"
/note="end : SP6"

BASE COUNT 254 a 176 c 160 g 152 t 359 others
ORIGIN
Query Match 5.3%: Score 63; DB 83; Length 1101;
Best Local Similarity 15.6%; Pred. No. 0.0045;
Matches 73; Conservative 206; Mismatches 189; Indels 0; Gaps 0;

```


Thu Jun 8 15:53:48 2000

us-08-900-220-8.rst

Page 11

THIS PAGE BLANK (u9PT0)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2000, 06:51:49 ; Search time 975.44 Seconds

(without alignments)
6739.859 Million cell updates/sec

Title: US-08-900-220-7

Perfect score: 1622
Sequence: 1 CATCAGCCACACAGAGACC.....CATGGGAGGCCCATTCCTCC 1622

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 9714632

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST: *
1: em_est1: *
2: em_est2: *
3: em_est3: *
4: em_est4: *
5: em_est5: *
6: em_est6: *
7: em_est7: *
8: em_est8: *
9: em_est9: *
10: em_est10: *
11: em_est11: *
12: em_est12: *
13: em_est13: *
14: em_est14: *
15: em_est15: *
16: em_est16: *
17: em_est17: *
18: em_est18: *
19: em_est19: *
20: gb_est1: *
21: gb_est2: *
22: gb_est3: *
23: gb_est4: *
24: gb_est5: *
25: gb_est6: *
26: gb_est7: *
27: gb_est8: *
28: gb_est9: *
29: gb_est10: *
30: gb_est11: *
31: gb_est12: *
32: gb_est13: *
33: gb_est14: *
34: gb_est15: *
35: gb_est16: *
36: gb_est17: *
37: gb_est18: *
38: gb_est19: *
39: gb_est20: *
40: gb_est21: *
41: gb_est22: *
42: gb_est23: *
43: gb_est24: *
44: gb_est25: *

45: gb_est26: *
46: gb_est27: *
47: gb_est28: *
48: gb_est29: *
49: gb_est30: *
50: gb_est31: *
51: gb_est32: *
52: em_est20: *
53: em_est21: *
54: em_est22: *
55: em_est23: *
56: em_est24: *
57: em_est25: *
58: em_est26: *
59: gb_est33: *
60: gb_est34: *
61: gb_est35: *
62: gb_est36: *
63: gb_est37: *
64: gb_est38: *
65: em_est27: *
66: em_est28: *
67: em_est29: *
68: em_est30: *
69: gb_est39: *
70: gb_est40: *
71: gb_est41: *
72: gb_est42: *
73: gb_est43: *
74: gb_est44: *
75: em_est31: *
76: em_est32: *
77: em_est33: *
78: em_est34: *
79: gb_est45: *
80: gb_est46: *
81: gb_est47: *
82: gb_gss1: *
83: gb_gss2: *
84: gb_gss3: *
85: gb_gss4: *
86: em_gss1: *
87: em_gss2: *
88: em_gss3: *
89: em_gss4: *
90: gb_gss5: *
91: gb_gss6: *
92: gb_gss7: *
93: gb_gss8: *
94: gb_gss9: *
95: em_gss5: *
96: em_gss6: *
97: em_gss7: *
98: em_gss8: *
99: em_gss9: *
100: em_gss10: *
101: em_gss11: *
102: gb_gss10: *
103: gb_gss11: *
104: em_gss12: *
105: gb_gss13: *
106: gb_gss14: *
107: gb_gss15: *
108: gb_gss16: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES

Db	502	GAAGGCTCTGG	512
RESULT	2		
LOCUS	AI1325370		
DEFINITION	AI1325370	452 bp	EST
ACCESSION	AI1325370		23-DEC-1998
VERSION	AI1325370.1		
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.		
TITLE	The Mashu-HMI Mouse EST Project		
JOURNAL	Unpublished (1996)		
COMMENT	On Jan 19, 1998 this sequence version replaced g1:2152292. Contact: Marra M/Mouse EST Project Mashu-HMI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810		
FEATURES			
SOURCE			
1..452	location/qualifiers		
/organism="Mus musculus"			
/strain="C57BL/6J"			
/db_xref="taxon:10090"			
/clone="IMAGE:465405"			
/clone_lib="Soares mouse embryo NBMEL3.5 14.5"			
/sex="unknown"			
/tissue_type="embryo"			
/dev_stage="13.5-14.5dpc total fetus"			
/lab_host="DH10B"			
/note="Vector: pRTT3D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer 15', TGTTACCAATCTGAAGTGGAGCGCGCGGAAATTTTTTTTTTTTTTTTTTTT 3'], on equal amounts of mRNA from 2 13.5dpc and 2 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne State Univ., from 2] ; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRTT3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Felina Bonaldo."			
BASE COUNT	70 a 137 c 140 g 105 t		
ORIGIN			
Query Match	19.9%; Score 323.2; DB 44; Length 452;		
Best Local Similarity	85.3%; Pred. No. 3.9e-54;		
Matches 372; Conservative	0; Mismatches 63; Indels 1; Gaps 1;		
960	GTGCTGCTGGCTGGAGTGCACGCTGCAGCCCGCGTGGACAGCTCTCTACACAC		
1019			

DB	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
DB	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
DB	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
DB	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
DB	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
DB	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
DB	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
DB	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
DB	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25																																			

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 291)	Bovinae; Bos.			
	Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Fairbanking,S.C., Fleking,B.A., Kohner,G.A. and Keefe,J.W.	Design and use of four pooled tissue normalized cDNA libraries for EST discovery in cattle	Unpublished (2000)	
		On Mar 10, 1998 this sequence version replaced gi:2948930.		
		Contact: Smith TPL		
		USDA, ARS, US Meat Animal Research Center		
		PO Box 166, Clay Center, NE 68933-0166, USA		
		Tel: 402 762 4366		
		Fax: 402 762 4390		
		Email: smiththe@mail.marc.usda.gov		
		Single pass sequencing. Bases called and trimmed with phred v0.980904.e. Vector identified by cross_match with the -mnscore 20 and -minmatch 12 options.		
		PCR primers		
		FORWARD: AGGAACAGCATGTACCAT		
		BACKWARD: GTTTTCCAGTCACGAC		
		Plate: 37 row: N column: 8		
		Seq primer: ATTGAGGACATATG.		
FEATURES	Source	Location/Qualifiers		
		1..291		
		/organism="Bos taurus"		
		/db_xref="taxon:9913"		
		/clone_lib="MARC 4BOV"		
		/issue_type="pooled"		
		/lab_host="DH10B"		
		/note="Vector: pcMV SPORT6; Site_1: XbaI; Site_2: XhoI; Library made from pooled tissue from day 20 and day 40 embryos."		
BASE COUNT	44 a	92 c	96 g	59 t
ORIGIN				
Query Match	14.3%;	Score 231.8;	DB 81;	Length 291;
Best Local Similarity	87.3%;	Pred.No.2.9e-36;		
Matches 254;	Conservative	0;	Mismatches 37;	Indels 0;
			Gaps	0;
QY 1079	GATGACATCTGCTTCGCGGGCGGTGCTGACCAACCATGCGCTCAGTTGGCCTTCGACC	1138		
Db 1	GATGACATCTGCTTCGCGGGCGGTGCTGACCAACCATGCGCTCAGTTGGCCTTCGACC	60		
QY 1139	CCTGAGACTCTTTCACAGCTTGCGATGGGGCAGCTGACCCCGGGGAGAGGTGCATTG	1198		
Db 61	CCTGAGACTCTTTCACAGCTTGCGATGGGGCAGCTGACCCCGGGGAGAGGTGCATTG	120		
QY 1199	GTAACCCAGAGCTCTACCGCGCTGGGGCGTCTCTGAGAGAGAGGAGGAGCTTCACACC	1258		
Db 121	GTAACCCAGAGCTCTACCGCGCTGGGGCGTCTCTGAGAGAGAGGAGGAGCTTCACACC	180		
QY 1259	ACTGGGCACTGTCGCGGGGAGGAGAGCTGAAAGAGACTCAACCGCTGCCTCTGGAAGTGT	1318		
Db 181	ACTGGGCACTGTCGCGGGGAGGAGAGCTGAAAGAGATCTTCTACTGCTCCCGCAAGACTGC	240		
QY 1319	GTAACGGGTCCAGAGAGCTCTACGACAGAGGAGGAGTGGCCTCTGGAAGGGA	1369		
Db 241	CACATGAGTCCAAAGGCGCTCCACACAGAGGAGGAGCTTGGCTTGGGAAGGA	291		
RESULT 8	AA628967/c	753 bp	EST	16-OCT-1997
LOCUS	af28967.s1	Scarses_total_fetus_NB2HF8_9w	Homo sapiens	cDNA clone
DEFINITION	IMAGE:1033019.3	mRNA sequence.		
ACCESSION	AA628967			
VERSION	AA628967.1	GI:2541354		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 753)	Eutheria; Primates; Catarrhini; Homidae; Homo.	Hillier, L., Allen, M., Bowles, L., Dubuque, T., Gelsel, G., Jost, S., Krizman, D., Kucaba, T., Lacey, N., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, R., Steptoe, M., Tan, F., Treisman, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.	WashU-NCI human EST project	Unpublished (1997)
				On Sep 12, 1996 this sequence version replaced g1:1407418.
				Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
				Email: est@wustl.wustl.edu This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information. Seq primer: -40m3 fwd. ET from Amersham High quality sequence stop: 474.
FEATURES	SOURCE			
		1..753		
		/organism="Homo sapiens"		
		/db_xref="taxon:9606"		
		/clone="IMAGE:1033019"		
		/clone_ld="Soares_Total-fetus_Nb2HF8_9w"		
		/dev_stage="8-9 weeks"		
		/lab_host="DH10B"		
		/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I - oligo(dT) primer [5' TGGTACCACTGAGTGGAGGAGGCGCGCTTAATTTTATTTT 3'] (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."		
BASE COUNT	190 a	199 c	219 g	145 t
ORIGIN				
Query Match	12.3%	Score 199.4	DB 36	Length 753
Best Local Similarity	96.8%	Pred. No. 8.4e-30		
Matches 214	Conservative 0	Mismatches 6	Indels 1	Gaps 1
QY 1402	CCTGCGCATGAAGATACACCATTTGAGACTGAGTGGGCAACACGAGTCCGCCACCG 1461			
Db 670	CCTTGGCCATAGATACACCATTTGAGACTGAGTGGGCAACACGAGTCCGCCACCG 611			
QY 1462	CGTGTGGTGTAGTCATGAGCTGCAAGCTAGCTGGGAGGGGATGGTTGACCCCT 1521			
Db 610	CGTGTGGTGTAGTCATGAGCTGCAAGCTAGCTGGGAGGGGATGGTTGACCCCT 551			
QY 1522	CTCTCTGAGAGACCTTGAGGCTGGACAGCGCACACTCCCACTCAGCTGCTCTCACTAGA 1581			
Db 550	CTCTCTGAGAGACCTTGAGGCTGGACAGCGCACACTCCCACTCAGCTGCTCTCACTAGA 491			
QY 1582	GTTTTCATATCTC-TGCTTCCCATTTGGGAGGGCCCATTC 1621			
Db 490	GTTTTCATATCTCGTCCCTCCCATTTGGGAGGGCCCATTC 450			
RESULT 9				
LOCUS	A1958076	406 bp	mRNA	EST
DEFINITION	fc90e01.y1 zebrafish washu mKING EST Danio rerio cDNA 5' similar to			
ACCESSION	TR:Q92008 Q92008 VHH-1 PRECURSOR ; mRNA sequence.			
VERSION	A1958076			
KEYWORDS	A1958076.1 GI:5750785			
SOURCE	EST.			
ORGANISM	zebrafish.			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;			

OY	195	TTCAGCAGATTACGCCCCCAATGTGCCCGGGAAGACCCTTGGGCCGACGCAGCTATGAA	254
DB	401	TACAAACAATTAGCCCCCAATGTGCCCGAAGAAGACCTTGGGCCGACGAGCTATGAA	342
OY	235	GGCAGATCGCTCGCAGCTCCGAGCGCTTCAAAGAGACTCACCCCATAATTAATCCAGAC	314
DB	341	GGCAGATCGCTCGCAGCTCCGAGCGCTTCAAATGAGCTCACCCCCATAATTAATCCAGAC	282
RESULT 11			
LOCUS	A1666359	463 bp	mRNA EST 12-MAY-1999
DEFINITION	mulig06.x1 Soares_thymus_2nbMT Mus musculus cDNA clone IMAGE:639130		
VERSION	3' similar to TR:Q61488 Q61488 DESERT HEDGEHOG HOMOLOG PRECURSOR ;		
KEYWORDS	mRNA sequence.		
SOURCE	A1666359	GI:4804713	
ORGANISM	EST. house mouse. Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 463) NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997)		
JOURNAL COMMENT	On May 18, 1998 this sequence version replaced gi:3136856. Other-ESIS: mulig06.y1 Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert.Strausberg@nih.gov This clone is available royalty-free through LUNL : contact the IMAGE Consortium (info@image.lnl.gov) for further information. This clone was previously sequenced on the 5' end only, this new data is from the 3' end Possible reversed clone: similarity on wrong strand High quality sequence stop: 261. Location/Qualifiers		
FEATURES			
source	1..463 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="IMAGE:639130" /clone_idb="Soares_thymus_2nbMT" /sex="male" /issue_type="Thymus" /dev_stage="4 weeks" /lab_host="DH10B" /note="Vector: pTR73D-Pac (Pharmacia) with a modified polylinker. Site.1: Not I; Site.2: Eco RI; 1st strand CDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGCAAGTGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTT 3'] : double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTR73 vector. RNA provided by Dr. Bertrand Jordan. Library went through two rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."		
BASE COUNT	87 a	124 c	138 g
ORIGIN	138 t	1 others	
Query Match	7.4%	Score 119.4:	DB 49; Length 463;
Best Local Similarity	68.0%;	Pred. NO. 3.6e-14;	
Matches 181;	Conservative	0; Mismatches 82;	Indels 3; Gaps 1;
OY	103	TGCTCTGCTGCTGCTGCCCCGCGCATCGGGGCTGGGCGCGGCTGGTGAGCGACCCGCC	162
DB	426	TGGCACTCTGTGCACATATCTGCCAGAGCTGGGGCGGCCGAGAACCGSGTGCCGAC	367
OY	163	GCGCAGCGCCACGCA---AATCTGTGCCGCTCGCTTACAAGCAGTTCAAGCCCAATGTGC	219

Db	366	GCCGTAATGTGGCAGACACTTGTGCCTTCGATATACAACAGATTGTGGCCAGTAGGC	307
Oy	220	CCGAGAATACTCTGGGCCCGCCAGCGACCTTTGAAGCAAAATGCTGCGAGCTCCGAGC	279
Db	306	CCGACCGGACCTTGCGCCCGAGTGGCCAGCGGAGGGAGGATTAACAAGGGGGTCCGAGC	247
Oy	280	GCTTCAGAGACTCATCCCCCAATTACATTCAGACATCATCTTCAAGAGCAGGAGAACA	339
Db	246	GCTTCGGGAGCTCGTACCACATCTCAACCCCAGCATATATCTTCAGAGATGAGGAGAACA	187
Oy	340	CAGGCGCGACCGCCTCATGACCCAG	365
Db	186	GCGGCGCAGACCGCCTGATGACAGAG	161
RESULT_12	A1645932		
LOCUS	A1645932	537 bp	mRNA EST 29-APR-1999
DEFINITION	multi96.v1 Soares_thymus_2NDBYT Mus musculus cDNA clone IMAGE:639130		
	5' similar to FR:061488 061488 DESERT HEDGEHOG HOMOLOG PRECURSOR ;,		
	mRNA sequence.		
ACCESSION	A1645932		
VERSION	A1645932.1	GI:4724407	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
AUTHORS	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
TITLE	Tumor Gene Index		
JOURNAL	Unpublished (1997)		
COMMENT	On May 18, 1998 this sequence version replaced gi:3137802. Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert.Strausberg@nih.gov This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. This read is a RESEQUENCE of a previously sequenced mouse clone This read has been verified (found to hit its original self in the correct orientation) Possible reversed clone: similarity on wrong strand Seq primer: -40RP from Gldco High quality sequence stop: 478. Location/Qualifiers 1..537 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="IMAGE:639130" /clone_id="Soares_thymus_2NDBYT" /sex="male" /tissue_type="Thymus" /dev_stage="4 weeks" /lab_host="DH10B" /note="Vector: pRT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTCACATCTGAAGTGGAGCGCGCCGCTTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT73 vector. RNA provided by Dr. Bertrand Jordan. Library went through two rounds of normalization, and was constructed by Bento Soares and M.Fatima Bernaldo."		
FEATURES	source		
BASE COUNT	107 a	159 g	107 t
ORIGIN			
Query_Match	7.1%	Score 115.6;	DB 49; Length 537;
Best Local Similarity	67.3%;	Pred. NO. 2.le-13;	

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)



Creation date: 02-12-2004

Indexing Officer: ATEKLY - ALEM TEKLAÿ

Team: OIPEScanning

Dossier: 08900220

Legal Date: 04-23-2001

No.	Doccode	Number of pages
1	SRNT	54

Total number of pages: 54

Remarks:

Order of re-scan issued on

